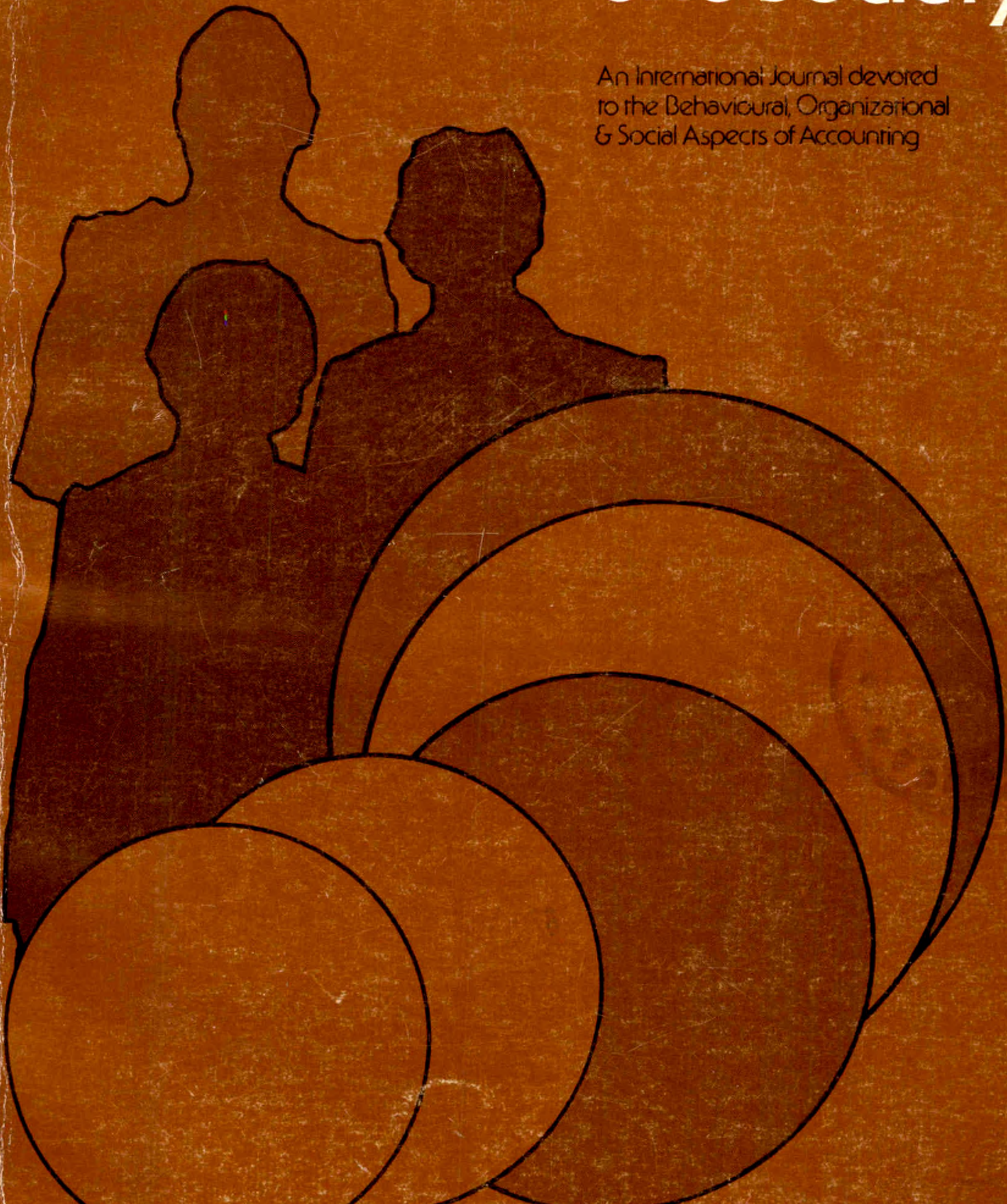
Volume 14 Number 4 1989

ISSN 0361–3682

Accounting, Organizations and Society

An International Journal devoted
to the Behavioural, Organizational
& Social Aspects of Accounting



Pergamon Press · Oxford · New York · Beijing · Frankfurt
São Paulo · Sydney · Tokyo · Toronto

Accounting, Organizations and Society

EDITOR-IN-CHIEF

Anthony G. Hopwood: London School of Economics and Political Science, Houghton Street
London WC2A 2AE

ASSOCIATE EDITORS

Jacob G. Birnberg
Graduate School of Business
University of Pittsburgh,
Pittsburgh

David Cooper
School of Management
University of Manchester
Institute of Science and
Technology

Peter Miller
Department of Accounting
and Finance
London School of Economics

EDITORIAL BOARD

Shahid L. Ansari
School of Business Administration and
Economics
California State University
Northridge

Chris Argyris
Graduate School of Business
Administration
Harvard University

Edward C. Arrington
College of Business Administration
University of Iowa

Stan Baiman
Graduate School of Industrial
Administration
Carnegie-Mellon University

Richard J. Boland
College of Commerce and
Business Administration
University of Illinois

Peter Brownell
School of Economics
and Financial Studies
Macquarie University

Edwin H. Caplan
Robert O. Anderson School of
Business and Administrative
Sciences
University of New Mexico

Jeremy Dent
Manchester Business School

Meinolf Dierkes
Science Center
West Berlin

Mark Dirsmith
College of Business Administration
Pennsylvania State University

Kenneth R. Ferris
Cox School of Business
Southern Methodist University

Eric G. Flamholtz
Graduate School of Management
University of California
Los Angeles

L. Richard Hoffman
Graduate School of Management
Rutgers University

Geert H. Hofstede
University of Limburg, Maastricht

Robin Hogarth
Graduate School of Business
University of Chicago

Trevor Hopper
Department of
Accounting and Finance
University of Manchester

Robert E. Jensen
Department of Business Administration
Trinity University

Sten Jönsson
Department of Business
Administration
Gothenburg University

Edward E. Lawler III
School of Business
Administration
University of Southern California

Robert Libby
Graduate School of Business
Administration
University of Michigan

James G. March
Graduate School of Business
Stanford University

Charles Medawar
Public Interest Research Centre
London

Kenneth A. Merchant
Graduate School of Business
Administration
Harvard University

John W. Meyer
Department of Sociology
Stanford University

Theodore J. Mock
School of Accounting
University of Southern California

Ted O'Leary
College of Commerce and Business
Administration
University of Illinois

David T. Otley
Department of
Accounting and Finance
University of Lancaster

Hein Schreuder
University of Limburg, Maastricht

Michael Shields
School of Accountancy
San Diego State University

Robert J. Swieringa
Financial Accounting Standards
Board

Sadao Takatera
Faculty of Economics
Kyoto University

William Waller
College of Business and
Public Administration
University of Arizona

John H. Waterhouse
Faculty of Business
Administration and Commerce
University of Alberta

Karl E. Weick
School of Business Administration
University of Michigan

Aaron Wildavsky
Survey Research Center
University of California
Berkeley

Mayer Zald
Department of Sociology
University of Michigan

Publishing, Subscription and Advertising Offices: Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW (*Oxford* 64881; *Telex* 83177).

Annual Subscription Rates 1989 (including postage and insurance)

Annual institutional subscription rate (1989) DM 730.00. 2 year institutional rate (1989/90) DM 1387.00. Personal subscription for those whose library subscribes at the regular rate (1989) DM 228.00. All subscription enquiries should be addressed to: The Subscription Fulfilment Manager, Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, U.K. Prices are subject to change without notice.

Back Issues: Back issues of all previously published volumes are available direct from Pergamon Press. Back issues of Pergamon journals in microform can be obtained from: UMI, 300 North Zeeb Road, Ann Arbor, MI 48106, U.S.A.

Published 6 times per annum

Copyright © 1989 Pergamon Press plc

It is a condition that manuscripts submitted to this journal have not been published and will not be simultaneously submitted or published elsewhere. By submitting a manuscript, the authors agree that the copyright for their article is transferred to the publisher if and when the article is accepted for publication. However, assignment of copyright is not required from authors who work for organizations which do not permit such assignment. The copyright covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microform or any other reproductions of similar nature and translations. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the copyright holder.

US Copyright Law Applicable to Users in the USA
Photocopying information for users in the U.S.A. The Item-Fee Code for this publication indicates that authorization to photocopy items for internal or personal use is granted by the copyright holder for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service provided the stated fee for copying beyond that permitted by Section 107 or 108 of the United States Copyright Law, is paid. The appropriate remittance of $3.00 per copy per article is paid directly to the Copyright Clearance Center Inc., 27 Congress Street, Salem, MA 01970.

Permission for other use. The copyright owner's consent does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific written permission must be obtained from the publisher for such copying.

The Item-Fee Code for this publication is: 0361-3682/89 $3.00 + .00

Colo-H02552-3-P006007

Accounting, Organizations and Society, Vol. 14, No. 4, pp. 303–319, 1989.
Printed in Great Britain

0361–3682/89 $3.00+.00
Pergamon Press plc

LIFO ADOPTION: A TECHNOLOGY DIFFUSION ANALYSIS*

BEN-HSIEN BAO
University of Colorado

and

DA-HSIEN BAO
Rutgers University

657.01
Ac27
V.14

③

Abstract

The results of prior research concerning the association between LIFO adoption decisions and their potential financial determinants are inconclusive. From the perspective of the theory of technology diffusion, the focus of these studies is limited to economic factors/forces from system outsiders. The effects of interactions among system members are ignored. Evidence from this study, which is based on a technology-diffusion approach, suggests that the LIFO adoption process is best explained partially by interactions with investors, creditors and public and partially by interactions with other firms. Future research may benefit from incorporating proxy variables for sociological factors such as imitation.

The association between LIFO adoption decisions and their potential determinants has been examined extensively in the accounting literature under a variety of hypotheses that focus on such factors as tax benefits, the agency effects of bonuses, the agency effects of debt covenants and the impact of political costs (Watts & Zimmerman, 1986). The test results of these hypotheses are mixed. One major reason may account for the inconclusiveness: only publicly available *financial* information is used to examine these hypotheses. Yet empirical evidence suggests that economic reasons alone may be insufficient to explain the choice among alternative accounting methods (Tritschler, 1970; Shank & Copeland, 1973) and that "factors other than those examined apparently have influenced the LIFO adoption decision" (Hunt, 1985, p. 466). The suggested non-economic factors include behavioral and sociological variables (Shank & Copeland, 1973; Hussein, 1981).

One approach to these variables comes from the theory of technology diffusion, which models the routes by which technologies are spread and adopted through a population. Several researchers have established the usefulness of this technology (innovation) diffusion theory to the study of accounting choice decisions (Tritschler, 1970; Comiskey & Groves, 1972; Hicks, 1978; Hussein, 1981). From a technology diffusion perspective, a firm may adopt LIFO for economic and/or sociological and behavioral reasons. The underlying drive may come from investors, creditors and public who are not directly in charge of the operation of the firm, yet have interest in, and may benefit from, the firm's economic fitness, i.e. adopting LIFO by the firm, the economic demand or welfare of the investors, creditors and public is considered. The drive may also come from interactions with other firms and here the focus is social fitness, not economic fitness, e.g. by adopting LIFO the firm's inventory policy is consistent with the "norm" of the industry to which it belongs.[1] However, none of the above-mentioned studies has specifically tested the process models well established in the technology diffusion literature. The purpose of this study is to examine the LIFO adoption decision by using non-financial data for testing three competing technology diffusion process models: the non-imitation model, the imitation model and the generalized model. The results should provide additional insights into the LIFO adoption process itself.

---

*The authors appreciate the helpful comments of A. Ali, W. R. Baber, B. Donabedian, C. Marais, V. Soybel, M. A. Vasarhelyi, two anonymous reviewers and the Editor.

The first section of this study presents a simplified framework for analyzing accounting diffusion processes. The next section is a summary of prior research. This is followed by a description of three basic diffusion process models. Later sections address the selection of the sample, the results of the tests and a discussion of the limitations and implications of the research.

A SIMPLIFIED FRAMEWORK

Technology diffusion can be broadly defined as the process by which information about a technology is communicated through certain channels over time among the members of a social system (Mahajan & Peterson, 1985). Accordingly, there are four major elements in the process: the technology, the channels of communication, time and the social system. The social system is defined as "a collectivity of units" that share a common objective (Rogers & Shoemaker, 1971).[2] The members are adopters and potential adopters of the technology. In line with the technology diffusion literature, they are labeled as system insiders. Those who are not directly in charge of the operations, such as the investors, creditors and public are system outsiders.[3] A technology is an idea, an object or a practice that is, in general, perceived as "new" by members of the social system. The technology need not be a recent invention (Shank & Copeland, 1973; Hussein, 1981); it may have been invented and known to members of the social system for some time, but the need for its adoption has not been felt by the non-adopters (Hussein, 1981). The channels of communication are the routes by which information is transmitted to or within the social system. They can be classified into two major categories: the vertical channels that communicate information *to* the social system and the horizontal channels that communicate *within* the social system. The element of time relates to the relative speed at which the technology is adopted.

A simplified framework for analyzing accounting technology diffusion is developed in accordance with the above descriptions. It is presented in Fig. 1. The technology is LIFO. Although a well-known inventory method, LIFO is "new" to the non-LIFO firms in the sense that the managers of these firms have not felt the need for adopting it.[4] The social system consists of accounting entities, i.e. firms; some are LIFO adopters, some are potential adopters. All of the firms have the objective of producing, selling products and/or rendering services to their customers and have to choose among inventory val-

---

[1] The theory behind economic fitness and social fitness will be discussed in the Technology Diffusion Process Models section.

[2] The system members can be individuals, business organizations, public organizations or nations depending upon the specific diffusion context. The members share a common objective (Rogers & Shoemaker, 1971) or a common culture (Mahajan & Peterson, 1985). In the LIFO adoption context, the social system consists of firms that have the objective of producing, selling products and/or rendering services to their customers and have to select an inventory method for external reporting purposes. All the COMPUSTAT firms fit this description.

[3] The investors are the owners of the firm while the managers are "hired" by the investors to operate the firm. The investors have the ownership while the managers have the control of the firm. Since the owners are not *directly* in charge of the decisions and operations of the firm, they are labeled as outsiders.

[4] Rogers (1962) states that "it really matters little, as far as human behavior is concerned, whether or not an idea is "objectively" new as measured by the amount of time elapsed since its first use or discovery. It is the newness of the idea to the individual that determines his reaction to it" (p. 13).

P, 6007

③

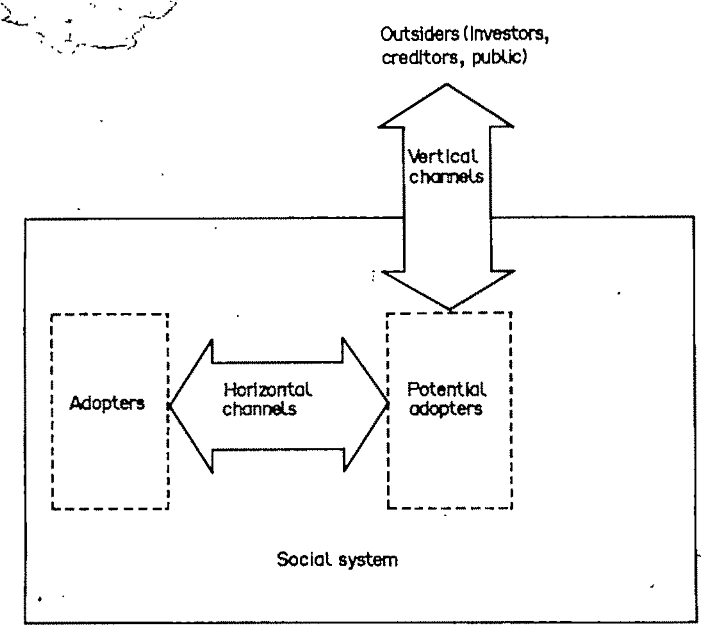


Fig. 1. A simplified framework of technology diffusion process.

uation methods for external reporting purposes. The interactions between the non-LIFO firms and investors, creditors and public and between the non-LIFO firms and LIFO firms may induce the potential adopters to adopt LIFO. The boundaries of the adopters and the potential adopters are not fixed, i.e. they are represented by dotted lines, so that the former may expand while the latter may shrink by passage of time. This framework is used to evaluate some of the prior research.

PRIOR RESEARCH

Prior research is classified into two groups: those studies focused on economics and those focused on sociological concepts. The economics studies focus only on the association between LIFO adoption decisions and their financial determinants. The sociologically informed inquirers attempt to identify the general non-economic factors that facilitate the diffusion of accounting technology.

Economic studies

Among what are here being classified as economic studies, four hypotheses have been proposed and tested to explain the choice between LIFO and other inventory methods. They are the tax-benefit hypothesis, the agency theory–bonus hypothesis, the agency theory–debt covenant hypothesis and the political-cost hypothesis. The tax-benefit hypothesis suggests that a manager chooses LIFO since his or her firm's net operating cash inflow is increased due to the reduction in tax liability. The manager of a firm may decide to stay on FIFO because a switch to LIFO may lead to a reduction of shareholders' wealth due to the lower reported income. Conceptually, tests of this hypothesis may be based on the association between LIFO adoption and cash flow position. Previous studies focus only on the association between LIFO adoption and its effect on the firm's shareholders' wealth. The shareholders' wealth is measured by the firm's security prices. It is hypothesized that by adopting LIFO, a firm's security price will be decreased because of the lower reported income. The results of tests are inconsistent. Abdel-

Khalik & McKeown (1978) and Ricks (1982) report a negative association between a LIFO switch and security prices. Brown (1980) reports no significant association, while Sunder (1973, 1975) and Biddle & Lindahl (1982) report a positive association.

The agency theory–bonus hypothesis focuses on the manager–shareholder relationship and assumes that shareholders of a firm will reward the manager if the firm's accounting earnings number is increased, i.e. shareholders' incentive plans influence the behavior of managers. This hypothesis proposes that a manager selects an income-increasing accounting method, e.g. FIFO, because his or her income-based bonus increases due to the choice. The results of these tests are also mixed. Hagerman & Zmijewski (1979) found that compensation was not a significant factor for explaining the choice of inventory method; even the sign of the regression coefficient was opposite to that suggested by the hypothesis. Abdel-Khalik (1985) found that changing to LIFO did not have a significant negative effect on compensation. Zmijewski & Hagerman (1981) report that the existence of a profit-sharing plan influences the manager to choose an income-increasing accounting strategy, e.g. the manager would choose FIFO. The accounting strategy of a firm is defined in their work as a combination of four choices. Each relates to the choice between accelerated and straight-line depreciation methods, LIFO and FIFO, amortization of past cost in less than 30 years and greater than or equal to 30 years and deferral and flow-through methods for investment tax credit, respectively.

The agency theory–debt covenant hypothesis focuses on the bondholder–shareholder relationship. It suggests that there is a positive relationship between a firm's contractual constraints and income-increasing accounting methods. Payment of dividends and incurrence of additional debt are constrained by the firm's lending agreements (Leftwich, 1983). A LIFO firm is more constrained than a FIFO firm because by choosing LIFO a firm has a lower reported income and a lower retained earnings. This will lead the firm to pay lower dividends because of the lower reported income. This will also make it more difficult for the firm to incur additional debt because of the higher debt–equity ratio. In general, payment of dividends is measured by the ratio of dividends to unrestricted retained earnings, while incurrence of additional debt is measured by the ratio of net tangible assets to long-term debt. The results of these tests are inconclusive. Morse & Richardson (1983) found that the two financial ratios were not statistically significant in explaining choice of inventory method, while Hunt (1985) found them to be significant.

The political-cost hypothesis suggests that a firm may avoid the attention caused by high reported profit and therefore choose LIFO, due to the public's association of high profit with monopoly rents (Watts & Zimmerman, 1978). Political cost is usually measured by the size, i.e. sales of the firm. The larger the sales of a firm, the higher the probability this firm will choose LIFO to report a lower income to avoid public attention. The results of these tests are also inconsistent. Hagerman & Zmijewski (1979) found that size is not statistically significant in explaining LIFO adoption while Morse & Richardson (1983) found just the opposite.

In line with the diffusion framework presented in Fig. 1, it may be concluded that these studies consider only part of the whole diffusion process. Their main focus, as represented in the framework, is on the effects of communication between firms and system outsiders. The managements of the potential adopters presumably base their decisions on the demands or welfare of the investors (tax-benefit hypothesis, agency theory–bonus hypothesis), the creditors (agency theory–debt covenant hypothesis) or the public (political-cost hypothesis). In this type of modeling, no consideration is explicitly given to the effects of communication among firms in the social system, i.e. no sociological factors such as imitation are considered.

Sociological studies

Several studies explicitly consider behavioral and sociological factors among the social system members and study the adoption distribution.

Tritschler (1970) used mainly aggregated non-economic data from published sources to examine the adoption pattern of LIFO and accelerated depreciation methods in the U.S. and price-level adjusted financial statements in France. His major result was that tax relief alone was not sufficient to explain the choice among alternative accounting methods. He conjectured, without tests, that the adoption process should be examined from a sociological perspective.

Copeland & Shank (1971) used a questionnaire methodology to examine LIFO adoption decisions. Their results supported Tritschler's (1970) view that profit maximization alone does not explain LIFO adoption decisions, but the weight assigned to behavioral factors was not significant. Shank & Copeland (1973) also used non-economic data from published sources to examine the behavior of "change-receptive" firms. They claimed that once a firm has learned the effects of one set of alternative accounting methods, it tends to adjust rapidly to the effects of other sets of alternative accounting methods, i.e. the firm has formed a learning set. Inventory, depreciation and investment tax credit methods were jointly considered. The results generally supported the "learning set" theory, which claims that a firm receptive to change shows more accounting changes in a given time period than a random sample of firms.

Rogers (1962) hypothesized that the adoption frequency distribution over time is a bell-shaped normal distribution. The cumulative adoption time path, i.e. a plotting of cumulative number of adopters against time, is an S-shaped curve. Comiskey & Groves (1972) used a questionnaire methodology to study the adoption distributions of their sample firms. Their focus was on the frequency distributions of innovations, rather than the process models. The results showed that the distributions were not normal. Any unimodal frequency distribution, however, will generate an S-shaped cumulative adoption time path. Comiskey & Groves' results may imply that the frequency distribution, although not normal, probably is still unimodal.

Hicks (1978) also used questionnaires to study the perceptions toward new accounting technologies by accounting interest groups. One major finding was that system effects are important factors in the diffusion process. System effects are "the influence of the norms, values, and functions of a group on the behavior of the individual members of that group" (Hicks, 1978, p. 373). There is a significant relationship between an individual's perception of a technology and his or her membership in an interest group. He finds that accounting academicians have a higher need for changes than CPAs and financial executives who belong to other interest groups. This concept can be used to highlight the importance of the communication between firms presented in Fig. 1. The other major finding was that new technologies often fail to gain adoption by the accounting profession because practitioners perceive the need for change differently from academicians. This may partially explain why many firms still have not adopted LIFO even though it is an existing and well-known technology.

Hussein (1981) studied the diffusion process for accounting standards using a questionnaire methodology. The questionnaires were sent to members of the FASB Advisory Council and of the FASB's sponsoring organizations to study the association between perception, communication channels and nine accounting standards. Hussein's major contribution is the establishment of a conceptual framework for analyzing this process. In Hussein's framework, he specifically recognized both the vertical and the horizontal channels of communication. The empirical results showed that financial accounting standard setting is indeed a two-dimensional diffusion process.

To date, most sociological oriented studies have used a questionnaire methodology to examine aspects of the accounting technology diffusion processes. Some support the diffusion theory and all recognize the importance of non-economic factors. Hussein (1981) acknowledged that "survey data are mostly useful in studies of a static nature" (p. 36) and suggested that other methodologies should be useful for studying processes. In addition, none of these studies specifically tests the three alternative process models proposed in the diffusion literature. The next section discusses these process models.

TECHNOLOGY DIFFUSION PROCESS MODELS

Three basic models have been developed to study the process of technology adoption. They are the non-imitation model (Fourt & Woodlock, 1960; Coleman, 1964), the imitation model (Dodd, 1958; Mansfield, 1968) and the generalized model (Rogers & Shoemaker, 1971; Mahajan & Schoeman, 1977). These three models serve as competing explanations of the diffusion process. The non-imitation model (also called the external-influence model) assumes that the sources of the diffusion process are external to the social system. There is no imitation, i.e. there is no interaction among system members. The imitation model (also called the internal-influence model) assumes that all adopters (except the first one, who is an innovator) are imitators, i.e. an adopter adopts a technology only after observing another adopter using the technology. The generalized model (also called the mixed-influence model) is a combination of the above two models. It assumes that the adoption decision is made partially through external sources and partially through imitation.

It is also useful to examine these models from a competitive or institutional isomorphism perspective (DiMaggio & Powell, 1983; Dirsmith, 1986). DiMaggio & Powell (1983) analyzed the homogeneity of organizational forms and practices. They maintained that there are two types of isomorphism: competitive and institutional. The competitive isomorphism concept emphasizes market competition. It suggests that the change in organizational forms and practices is driven by competition and the need for efficiency and economic fitness; social fitness is not considered. This concept is consistent with the non-imitation model. DiMaggio & Powell (1983), however, stated that competitive isomorphism "may apply to early adoption of innovation, but does not present a fully adequate picture of the modern world of organizations" (p. 150). They suggested that organizations aggressively compete for political power and institutional legitimacy, in addition to resources and customers, therefore institutional isomorphism is introduced as an additionally useful concept.

The institutional isomorphism concept emphasizes social fitness, i.e. organizations tend to become isomorphic despite the differences in underlying technical operations. Institutional isomorphic change occurs through three mechanisms: coercive, mimetic and normative. Coercive isomorphism results from external constituents' pressures, mainly from government mandate, to adopt new practices. Mimetic isomorphism results from organizational copying, i.e. organizations model themselves after their peers that are perceived as more successful. Normative isomorphism results from professionalization, i.e. the collective struggle of the members of a profession to establish the profession's legitimacy and to define the conditions and methods of its members' works. It is achieved through formal education and professional networks. The imitation model discussed before is generally supported by the concepts of mimetic and normative isomorphism. The generalized model basically fits the descriptions of both competitive and institutional isomorphism.

The study of LIFO adoption decision in the technology diffusion framework has been suggested by Copeland & Shank (1971) and Tritschler (1970). In the LIFO adoption decision context, the LIFO method is the technology. The non-imitation model claims that the adoption of LIFO is induced (or forced) mainly by interaction with the investors, creditors and public. The manager of a non-LIFO firm may adopt LIFO for economic fitness reasons, for example, he or she feels pressure from the investors to improve the firm's cash flow position. Social fitness, e.g. other firms' selection of inventory methods, has no influence on the decision. The imitation model claims that the diffusion process is mainly caused by imitation; therefore, a firm adopts LIFO with the rationality of mimetic or normative isomorphism. The manager of a non-LIFO firm may adopt LIFO simply because other firms in the social system choose LIFO. The demand or welfare of the investors, creditors and the public is not the major concern. The generalized model claims that the diffusion process can be explained partially by external factors and partially by imitation, i.e. with the rationality of competitive and institutional isomorphism.

If the number (population) of potential adopters is fixed,[5] then the three models can be specified as follows:

(1) non-imitation model:

$$\frac{dk_t}{dt} = \text{beta}_1\,(N - k_t), \qquad (1)$$

(2) imitation model:

$$\frac{dk_t}{dt} = \text{beta}_2 k_t\,(\frac{N-k_t}{N}) \text{ and} \qquad (2)$$

(3) generalized model:

$$\frac{dk_t}{d_t} = \text{beta}_1\,(N - k_t) + \text{beta}_2 k_t\,(\frac{N-k_t}{N}), \qquad (3)$$

where N is the number of potential adopters, k_t is the cumulative number of adopters at time t,[6] beta_1 is an index of uninfluenced adoption and beta_2 is an index of influenced adoption (by imitation).[7]

The non-imitation model assumes that the diffusion process proceeds from an outside source, independent of the number of firms that have already adopted LIFO. The impact of the external source on non-LIFO firms is represented by beta_1. The number of new LIFO adopters per unit time is proportional to the total number of firms that have not yet adopted LIFO. The imitation model assumes that firms are able to imitate. The LIFO adoption rate is dependent not only on the number of firms that have adopted LIFO but also on the proportion of the maximum number of potential adopters which are still using a non-LIFO method. The term $k_t\,(N - k_t)$ represents interaction between the adopters and the potential adopters at time t. The generalized model is a combination of the above two models. The solutions for equations (1), (2) and (3) are:

(1) non-imitation model:

$$k_t = N - \exp(\ln(N - k_0) - \text{beta}_1\,(t - t_0)), \qquad (4)$$

(2) imitation model:

$$k_t = \frac{N\exp(\ln(\frac{k_0}{N-k_0}) + \text{beta}_2\,(t - t_0))}{\exp(\ln(\frac{k_0}{N-k_0}) + \text{beta}_2\,(t - t_0)) + 1} \text{ and} \qquad (5)$$

(3) generalized model: (6)

$$k_t = \frac{N\exp(\ln(\frac{k_0\text{beta}_2 + N\text{beta}_1}{N-k_0}) + (t-t_0)(\text{beta}_1 + \text{beta}_2)) - \text{beta}_1 N}{\exp(\ln(\frac{k_0\text{beta}_2 + N\text{beta}_1}{N-k_0}) + (t-t_0)(\text{beta}_1 + \text{beta}_2)) + \text{beta}_2},$$

where k_0 is the cumulative number of adopters at time t_0.

The cumulative adoption time paths (k_t) generated by the three models, although S-shaped, are not identical because of the assumptions made, i.e. adoption is caused by interacting only with system outsiders, only with other firms or both. A hypothetical example is provided in Fig. 2 to highlight the behavioral patterns. In this example, the number of adopters at t_0 is one, the total number of potential adopters is 100, the impact of uninfluenced adoption (beta_1) is 0.01 while the impact of influenced adoption (beta_2) is 0.2. Figure 2 shows the 20-year patterns of

---

[5] The process models are fixed population models. One requirement is that the number of potential adopters, N, is fixed and known. Although sample firms are drawn from publicly-held companies listed in the COMPUSTAT tape, some publicly-held firms are not listed and their inventory methods are therefore not known. Some of the listed firms' inventory methods are not disclosed on the tape. Given these constraints, this study is more exploratory than conclusive. Other limitations are also discussed in the Conclusion section.

[6] When applying the three basic process models, a global assumption is that "all relevant information has been captured by the models", usually through the term k_t (Mahajan & Peterson, 1985, p. 25). This is perhaps a too-strong assumption. However, these models are basic and can be used as the starting points for studying diffusion processes.

[7] The nature of the indices is very similar to that of the regression coefficients of a regression equation. The values of the regression coefficients can be negative, zero or positive. The values of the two indices in the technology diffusion models, however, are either zero or positive. The reason is that the models consider only the adoption process, not the abandonment process, i.e. the *cumulative* number of adopters can only stay the same or increase over time.

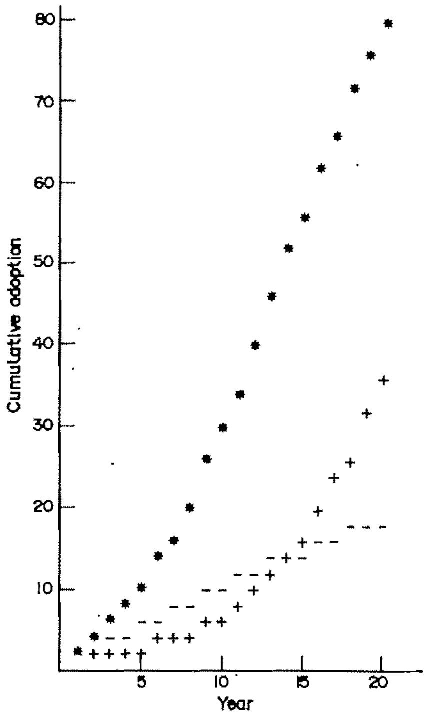


Fig. 2. An example. – Model 1, + model 2, * model 3.

adoption for the non-imitation model (model 1), the imitation model (model 2) and the generalized model (model 3). The adoption patterns are different by visual inspection.

Coleman (1964) has illustrated the usefulness of these idealized models by investigating the process of the adoption of a drug by doctors in four midwestern communities. He has found that the non-imitation model best describes the adoption behavior of the "isolated" doctors, while the imitation model best describes the adoption behavior of the "integrated" doctors. The important factor is a doctor's integration or isolation in the local medical community, as measured by the number of times other doctors named him or her as a social friend or as a frequent fellow-discussant of medical cases.[8] The "isolated" doctors depend on advertising, the drug salesmen and occasionally informed patients (outsiders) to make their decisions. The "integrated" doctors are in the social position to discuss with and learn from other doctors (insiders) and adopt the drug. In general, the non-imitation model is appropriate when system members are "isolated" or when information about the technology is only available from an external source. The imitation model is appropriate "when a technology is complex and socially viable, not adopting it places system members at a comparative disadvantage" (Mahajan & Peterson, 1985, p.18).

The LIFO adoption process and the drug adoption process are similar in many respects. Each has a technology (LIFO vs drug), social system insiders (other firms vs other doctors), social system outsiders (investors, creditors, public vs advertising, drug salesmen, patients), interactions between non-adopters and adopters and between non-adopters and outsiders through communication channels and adoption rates. In the LIFO adoption context, the focus, however, is not on the degree of integration or isolation, i.e. it is not on the behavioral difference between the "isolated" firms and the "integrated" firms since the degree of integration or isolation of a firm is not easily measurable and both types of firms are potential adopters. The focus is on the main determinants or drives of the adoption process of *all* the potential adopters, i.e. whether the managers of the non-LIFO firms make their adoption decisions by mainly considering the demand or welfare of and being influenced by the outsiders, by mainly considering the inventory method decisions of the adopting firms or by considering both.

Realistically, the non-imitation model and the imitation model are limited and may not be entirely satisfactory for explaining accounting method choice by *all* the non-adopters while the generalized model is more appropriate. There are two plausible reasons for the generalized model's appropriateness. Following Coleman's (1964) rationale, the non-adopters are either "isolated" firms or "integrated" firms and their adoption process can best be described by the non-imitation model and the imitation model, respectively. The adoption process of the combined group, i.e. *all* the non-LIFO firms, therefore probably can only be described by the generalized model. The other reason is that due to the advances in communication, the manager of a firm does have more opportunities to discuss the firm's decisions with outsiders, such as with the investors through shareholders' meetings and to learn the advantages of adopting a specific accounting method through professional association meetings and publications and personal contacts. The second rationale is probably closer to the reality and more consistent with DiMaggio & Powell's (1983) isomorphism concepts. But even with the opportunities of interacting with both the outsiders and the insiders, the manager of a non-LIFO firm may still make his or her LIFO adoption decision for the firm by mainly considering the outsiders' demand or welfare (the non-imitation model), by mainly considering other firms' choices (the imitation model) or by considering both (the generalized model).

---

[8] According to Coleman (1964), isolation does not necessarily imply physical isolation. The important element is whether a doctor maintains communication with other doctors.

Accordingly, the following competing hypotheses are made.

> *Hypothesis 1.*
> Non-LIFO firms adopt LIFO mainly after their managers' consideration of the demand or welfare of the system outsiders (investors, creditors and public). Although the managers have opportunities to communicate with and learn from the managers of LIFO firms, this type of interaction has no significant influence on the LIFO adoption process.

> *Hypothesis 2.*
> Non-LIFO firms adopt LIFO mainly after their managers communicate with and learn from the managers of LIFO firms. Although the managers have opportunities to consider the demand or welfare of the system outsiders (investors, creditors and public), this type of interaction has no significant influence on the LIFO adoption decision.

> *Hypothesis 3.*
> Non-LIFO firms adopt LIFO after their managers consider the demand or welfare of the system outsiders (investors, creditors and public) and after their managers communicate with and learn from the managers of LIFO firms.

As stated earlier, hypotheses 1 and 2 are restrictive and expected to be rejected while hypothesis 3 is more realistic in the LIFO context and is not expected to be rejected.

SAMPLE AND ANALYSES

In order to examine the LIFO adoption process in the 1965–1984 period, sample firms were selected from the COMPUSTAT tape.[9] The "split-half" approach was adopted and four accuracy-of-forecast tests were performed to compare the process models.

Sample and 1965–1984 analyses

A firm was included in the sample if both of the following criteria were met.

(1) the firm existed during the entire 1965–1984 period; those firms that merged, bankrupted or were newly incorporated after 1965 are excluded. Merger is defined as a major merger.[10]

(2) The COMPUSTAT data item number 59 clearly identifies the firm's inventory valuation

---

[9] For accounting purpose, a firm reports the use of LIFO under two situations: (1) it chooses LIFO for tax purpose; it then must choose LIFO for accounting reporting purpose under the conformity rule of the IRS; and (2) it chooses non-LIFO for tax purpose but decides to choose LIFO for accounting reporting purpose. The former may be explained by the tax-benefit hypothesis, i.e. the adoption decision is influenced only by a tax factor. The latter probably is influenced by interactions with system members and by external factors other than tax. Both situations can be explained by the technology diffusion theory. The focus of this study is the LIFO adoption process for accounting reporting purpose and the samples are selected from the COMPUSTAT tape. The inventory methods chosen by firms for tax purposes are not available. Even if they are known and if most of the firms adopt LIFO for tax reasons (therefore using LIFO for accounting reporting), the results should be "biased" in favor of hypothesis 1, i.e. the decision is influenced by outsiders such as investors and the IRS. The results of this study, however, suggest that non-economic factors are also important, i.e. hypothesis 1 is rejected while hypothesis 3 is not.

[10] A major merger is identified by COMPUSTAT footnote item number 1 code AB and its related codes.

TABLE 1. Cumulative number of LIFO adopters

Year	Observed
1965	8
1966	8
1967	8
1968	8
1969	8
1970	8
1971	10
1972	10
1973	12
1974	31
1975	33
1976	37
1977	39
1978	46
1979	56
1980	65
1981	70
1982	72
1983	75
1984	82

Sample size = 623.

method from 1965 to 1984.[11]

A firm in the sample was classified as a LIFO firm in a particular year if all its inventories were valued on the LIFO basis in that year.[12] However, this firm was not classified as a LIFO firm if it switched to other inventory methods in a later year.[13] The size of the sample is 623 and the *cumulative* numbers of LIFO firms for each of the years are reported in Table 1.[14]

The "split-half" approach requires the time period to be divided into two subperiods. The first is an estimation period while the second is a forecasting period. The estimation period in this section is from 1965 to 1977 and the forecasting period is from 1978 to 1984.[15]

The 1965–1977 data are used to estimate the parameter values for the non-imitation model (model 1), the imitation model (model 2) and the generalized model (model 3). Although equations (4) and (5) for models 1 and 2 are non-linear equations, linear transformations can be done and beta$_1$ and beta$_2$ are estimated by a linear regression procedure. Equation (6) for model 3 is also a non-linear equation but here no linear transformation is possible. The parameters are estimated by the hill-climbing procedure described in Kalwani & Morrison (1977). The estimated parameter values are reported in Table 2. t- and F-statistics and R^2 values for models 1 and 2 are also presented.[16] The re-

---

[11] The sample size is reduced to 2300 by criterion (1). It is further reduced to 623 by criterion (2) mainly because many firms' inventory methods prior to 1973 are not disclosed on the tape. The final sample size is still significantly larger than those in previous studies.

[12] There are two other LIFO classification schemes: "any LIFO" and "predominantly LIFO". "Any LIFO" means that a firm is classified as a LIFO firm in a particular year even if only a small portion of its inventory is valued on the LIFO basis in that year. This scheme is not adopted since a single-digit code was used in COMPUSTAT for years prior to 1972. A major classification problem can be illustrated by the following example: in 1971, a firm used FIFO for product A, which had 80% of the total inventory value and LIFO for product B. The single-digit code did not disclose the use of LIFO for product B. "Predominantly LIFO" means that a firm is classified as a LIFO firm in a particular year if the majority of its inventory is valued on the LIFO basis, i.e. the first digit of the code is a "2", in that year. The major problem is illustrated as follows: a firm used LIFO for product A and FIFO for product B in 1972 and 1973. It reported 55% LIFO and 45% FIFO at the end of 1972. During 1973 purchases for both products were the same but product A was sold at a faster rate. At the end of 1973 the firm reported 40% LIFO and 60% FIFO. The firm is classified as a LIFO firm only for 1972 even though the firm used inventory methods "consistently". The classification scheme used in this study is restrictive and is not perfect, but should be more reliable than the alternatives.

[13] Since the described diffusion models do not specifically deal with "switch-backs", those firms are not classified as LIFO firms.

[14] The table shows the *cumulative* number of adopters in each year, i.e. the number of firms in the sample, up to that year, adopted LIFO for all of its inventories. The numbers seem to be low because the "all-LIFO" criterion was used due to the COMPUSTAT coding problem discussed before.

[15] The 1965–1984 period is divided into two subperiods, the first twice as long as the second. This provides a longer estimation period and hence more accurate estimates, while the forecasting period remains long enough for calculating forecasting statistics.

[16] All the R^2 values reported are adjusted R^2 values.

TABLE 2. Estimated parameter values (1965–1977)

Parameter	Model 1	Model 2	Model 3
Beta_1	0.004662 ($t = 5.51$)*		0.000001
Beta_2		0.158113 ($t = 6.29$)*	0.13
F	30.31*	39.57*	
R^2	0.73	0.78	

* Significant at alpha = 0.01.

gression models are useful since the t- and F-statistics are significant and R^2 values are high. No regression statistics are reported for model 3 since the coefficients are not estimated through a regression procedure.

The estimated parameter values reported in Table 2 are used in equations (4), (5) and (6) to generate the predicted number of LIFO firms in the 1978–1984 period, i.e. the forecasting period. The predicted numbers are reported at the top of Table 3. The closeness of the prediction (to actual observation) is examined by four accuracy-of-forecast tests: the chi-square value, the mean-square error, the regression measures and Theil's U (Maddala, 1977).[17] The chi-square value tests the closeness between the forecasted and the actual numbers of LIFO firms. An insignificant chi-square value means that the prediction of that model is accurate. The mean-square error test is selected since it "seems to be the most popular" among mean-square error, mean absolute error and minimum number of errors in the prediction of turning points (Kennedy, 1979, p. 149). The model with the smallest

TABLE 3. Accuracy of forecast tests (1978–1984)

		Predicted		
Year	Observed	Model 1	Model 2	Model 3
1978	46	44.17	57.46	41.03
1979	56	46.86	66.25	46.30
1980	65	49.54	76.22	52.19
1981	70	52.21	87.44	58.75
1982	72	54.86	100.01	66.04
1983	75	57.50	114.01	74.12
1984	82	60.13	129.48	83.04
Chi-square (d.f. = 4)		31.38	47.61	8.49*
Mean-square error		0.067	0.169	0.022
Regression measures:				
intercept		−40.656	26.580	20.629
regression coefficient		2.055 ($t = 9.115$)	0.444 ($t = 6.683$)	0.763 ($t = 6.817$)
Theil's U		0.0036	0.0058	0.0021

* Not significant at alpha = 0.01.

---

[17] The chi-square value should be used as the primary criterion to judge forecasting accuracy because its significance can be tested. The other three tests are supplementary. There is no statistical test of significance for the mean-square error or Theil's U. The results of the regression measures are weak in this study.

mean-square error is the best model (Maddala, 1977, p. 344). A regression measure test that regresses actual values on predicted values is suggested by Mincer & Zarnowitz (1969). This is also a "popular" method and "if the intercept estimate tests insignificantly different from zero and the slope coefficient tests insignificantly different from one" the model is said to be a good one (Kennedy, 1979, p. 156). Theil's U is another popular test (Maddala, 1977, p. 346). It is zero in the case of perfect forecasts. The model with the smallest U is classified as the best model.

The chi-square values for comparing actual and predicted numbers of LIFO firms in the 1978–1984 period are presented in Table 3. Of the three models, only the third predicted numbers of LIFO firms that were not significantly different from the observed numbers. On this count, model 3 is the best model. Furthermore, Table 3 indicates that model 3 has the smallest mean-square error and Theil's U statistic and is, therefore, the best model according to these criteria as well. The intercepts of regression measures for all models are different from zero. The deviation of model 3's regression coefficient from one, however, is the smallest, i.e. 24% vs 106% for model 1 and 56% for model 2. The results given in Table 3 imply that the generalized model best explains LIFO-adoption decisions; hypotheses 1 and 2 are rejected while hypothesis 3 is not.

Subperiod analyses

Different firms may adopt LIFO in different years for a variety of reasons, therefore, the validity of the three diffusion process models can only be evaluated for a particular time period. According to the results of the last section, the LIFO adoption process in the 1965–1984 time period conforms most closely to model 3. If this time period is subdivided, then the validity of these three models has to be re-evaluated in accordance with the shorter time periods.

The American Institute of Certified Public Accountants (1975) and Abdel-Khalik (1985) reported a significant increase in LIFO firms in 1974. It is possible that the LIFO adoption process in the 1965–1973 period is entirely different from that of the 1974–1984 period, i.e. LIFO adoption decisions may be induced by different forces during the two time periods. This can be stated in the null form of a hypothesis.

Hypothesis 4.
LIFO adoption process in the 1965–1973 period is the same as the process in the 1974–1984 period.

The sample is divided into two periods: 1965–1973 and 1974–1984. The 1965–1973 data and equations (4), (5) and (6) are respectively combined to estimate parameter values. The estimated values are reported in Table 4. The available regression statistics for models 1 and 2 indicate that the estimation results are useful, i.e. t- and F-statistics are significant while R^2 values are high.

The estimated parameter values then are used to forecast the number of LIFO firms in the 1974–1984 period. The accuracy tests are also performed and the results are presented in Table 5. The chi-square values are significant. The mean-square errors, regression measures and

TABLE 4. Estimated parameter values (1965–1973)

Parameter	Model 1	Model 2	Model 3
Beta$_1$	0.000706 (t = 3.74)*		0.0001
Beta$_2$		0.046333 (t = 3.87)*	0.03
F	14.01*	14.94*	
R^2	0.67	0.68	

* Significant at alpha = 0.01.

TABLE 5. Accuracy of forecast tests (1974–1984)

Year	Observed	Predicted Model 1	Model 2	Model 3
1974	31	11.90	12.06	11.07
1975	33	12.33	12.62	11.46
1976	37	12.76	13.21	11.87
1977	39	13.19	13.82	12.28
1978	46	13.62	14.46	12.71
1979	56	14.05	15.13	13.15
1980	65	14.48	15.83	13.61
1981	70	14.91	16.56	14.07
1982	72	15.34	17.32	14.55
1983	75	15.77	18.12	15.05
1984	82	16.20	18.95	15.56
Chi-square (d.f. = 8)		1442.67	1217.04	1580.26
Mean-square error		0.663	0.625	0.691
Regression measures:				
intercept		−125.735	−67.520	−107.94
regression coefficient		12.869 (t = 18.029)	8.026 (t = 17.161)	12.34 (t = 17.760)
Theil's U		0.0136	0.0132	0.0139

Theil's U's all suggest that the adoption processes in the two periods are significantly different. However, an outright rejection of hypothesis 4 is inappropriate. The subdivision of the 1965–1984 period into the estimation and forecasting periods is "forced" by the observation that there is a significant increase in LIFO firms in 1974. The estimation period, therefore, is shorter than the forecasting period; the length of the estimation period is considerably shorter than twice of that of the forecasting period. This may have effects on the accuracy of parameter estimates and subsequently bias the forecasting results.

An alternative and indirect approach for testing hypothesis 4 is the independent examinations of the subperiods. The rationale is that if the data from the subperiods are supported by different technology diffusion models then the adoption processes in these subperiods are different. The conclusion is inconsistent with the previous results and the validity of the previous results should be questioned. If the data from the two subperiods are supported by the generalized model, i.e. the best model for the 1965–1984 period, then the explanatory factors are consistent in these subperiods. However, if the data from the two subperiods are supported by a technology diffusion model which is not the

TABLE 6. Estimated parameter values (1965–1969)

Parameter	Model 1	Model 2	Model 3
Beta$_1$	0 (t = 1.00)		0.00001
Beta$_2$		0.0000* (t = 1.00)	0.00013
F	0	0	
R^2	1	1	

* A very small positive number.

TABLE 7. Accuracy of forecast tests (1970–1973)

Year	Observed	Predicted		
		Model 1	Model 2	Model 3
1970	8	8	8	8.04
1971	10	8	8	8.04
1972	10	8	8	8.05
1973	12	8	8	8.06
Chi-square (d.f. = 1)		3*	3*	2.87*
Mean-square error		0.088	0.088	0.084
Regression measures:				
intercept		−1.92E16	−6.39E16	−1.33E3
regression coefficient		2.40E15	7.99E15	1.67E2
		($t = 2$)	($t = 4.270$)	($t = 39.38$)
Theil's U		0.0276	0.0276	0.0271

* Not significant at alpha = 0.01.

generalized model then the test results are inconsistent and useless.

The 1965–1973 data are divided into the two subperiods, 1965–1969 and 1970–1973. The 1965–1969 data are used to estimate parameter values, reported in Table 6, for forecasting the number of LIFO firms in the 1970–1973 period.[18] Table 7 summarized the results of accuracy-of-forecast tests. It shows that any one of the models predicts well, i.e. a non-significant chi-square value, a small mean-square error and a small Thiel's U. However, the results of accuracy tests may be of questionable value since the t- and F-statistics from the estimation period for models 1 and 2, reported in Table 6, are not significant. Nonetheless, the generalized model is the best model since its statistics from the forecasting period are the smallest among the three models.

The 1974–1984 data are also divided into two subperiods: 1974–1979 and 1980–1984. The 1974–1979 data are used to estimate parameter values. The results, presented in Table 8, are used to predict the number of LIFO firms in the

TABLE 8. Estimated parameter values (1974–1979)

Parameter	Model 1	Model 2	Model 3
Beta$_1$	0.008171 ($t = 6.33$)*		0.0090
Beta$_2$		0.122625 ($t = 8.77$)*	0.0002
F	40.09*	76.99*	
R^2	0.91	0.95	

* Significant at alpha = 0.01.

---

[18] In order to perform the chi-square test, a degree of freedom (d.f.) of at least one should be obtained. This causes the estimation period to be, although longer, less than twice of the forecasting period. The results, therefore, are weaker than those for the entire 1965–1984 period.

TABLE 9. Accuracy of forecast tests (1980–1984)

Year	Observed	Predicted Model 1	Predicted Model 2	Predicted Model 3
1980	65	59.32	61.38	62.17
1981	70	63.91	68.51	67.21
1982	72	68.46	76.35	72.19
1983	75	72.97	84.95	77.15
1984	82	77.45	94.35	82.05
Chi-square (d.f. = 2)		1.629*	3.277*	0.305*
Mean-square error		0.004	0.013	0.000
Regression measures:				
intercept		13.943	36.189	16.215
regression coefficient		0.860 ($t = 8.150$)	0.475 ($t = 9.325$)	0.784 ($t = 0.096$)
Theil's U		0.0008	0.0015	0.0003

* Not significant at alpha = 0.01.

1980–1984 period. Accuracy-of-forecast tests, reported in Table 9, show that the generalized model predicts best. Based on Tables 7 and 9, both the 1965–1973 LIFO adoption process and the 1974–1984 process can also best be explained by a generalized model. This confirms the results from the 1965–1984 process: the LIFO adoption process can best be represented by considering the effects of communication both with outsiders and with other firms that have already adopted LIFO.

CONCLUSION

The findings of this research should be considered within the content of its limitations. The first limitation is that the specified diffusion models are fixed-population models. The sample of this study is drawn from the COMPUSTAT tape which includes firms that trade their stocks on the major exchanges. The firms included in any one year are different from the firms included in another year. The reasons for exclusion are numerous. If a firm is not included in the tape for any one year then it is deleted. Bankrupted and newly incorporated firms, therefore, cannot be included in the sample and the dynamic aspect of growth or decline of the population is ignored. This problem limits the scope of this study. Sharif & Ramanathan (1984) have developed several dynamic models by changing the variable N in equations (1), (2) and (3) to N_t, i.e. the population is a prespecified and well-behaved algebraic function of time. Specification is difficult and mathematical tractability is low even if a well-behaved function can be found.

The second limitation is that the diffusion models focus only on the adoption process, ignoring the "abandonment" process. A firm may decide to switch from LIFO to other inventory methods. This problem may not be serious since Lindahl (1986) claimed that the number of "switch-backs" is considerably smaller than the number of LIFO adopters.

Despite these limitations, the results of this study do indicate that: (1) the LIFO adoption decision process in the 1965–1984 period can best be explained partially by external sources and partially by imitation; and (2) the adoption processes in the 1965–1973 period and in the 1974–1984 period can also best be explained by the joint influence of external factors and imitation; the explanatory factors for the two subperiods are the same although the "weighting" of the factors may differ. During the period as a whole, LIFO adoption decisions were made based on the demand or welfare of the investors, creditors and public, i.e. after the economic consequences to the outsiders were evaluated and on the managers' communication and professional networks among themselves. This study empirically establishes the importance of non-economic factors for explaining LIFO adoption from a technology diffusion perspective.

Several general observations can be offered to conclude this study. Numerous studies have been done to explain the choice of accounting methods by firms. Some focused on the diversity of choices, i.e. why some firms chose a method while other firms chose another method. Others focused on the similarity of choices, i.e. why many firms chose the same accounting method for reporting purposes. Most of the studies, however, focused on accounting choices *per se*. No attention was paid to the organizational and social environment in which accounting functions.

Accounting does not exist in vacuum. It is a tool used by firms to plan and control their operations and to communicate the results of operations. It is a vital component of organizations with various structures and behaviors. Although firms may differ in structure and behavior, DiMaggio & Powell (1983) argued that there is also a tendency towards homogeneity in organizational forms and practices because of connectedness and structual equivalence. i.e. firms are tied to each other through transactions or network structures. The homogeneity in practices, e.g. accounting choices, may be driven by the reasons of economic fitness, i.e. the economic fitness of the firm can be improved and the need of its constituents can be fulfilled if a firm adopts a particular accounting method. The homogeneity may also be driven by the reasons of social fitness, i.e. a firm is more socially fit if it adopts an accounting method which is used by a more successful firm or the professional managers choose the same accounting method to legitimize their occupational autonomy.

The economic fitness argument has been tested by solely examining the relationship between financial data and accounting choices. The results are inconclusive mainly because the social fitness argument was entirely ignored. The major argument of this study is that in order to understand accounting choices in general, and the LIFO adoption decision in particular, both the economic and the social reasons should be considered. Evidence based on non-financial data and technology diffusion models basically supports this argument. There is much to be gained if future research on accounting choices can take the sociological factors into account.

BIBLIOGRAPHY

Abdel-Khalik, A. R., The Effect of LIFO-switching and Firm Ownership on Executives' Pay, *Journal of Accounting Research* (Autumh 1985) pp. 427–447.

Abdel-Khalik, A. R. & McKeown, J. C., Understanding Accounting Changes in an Efficient Market: Evidence of Differential Reaction, *The Accounting Review* (October 1978) pp. 851–868.

American Institute of Certified Public Accountants, *Accounting Trends and Techniques* (New York: AICPA, 1975).

Biddle, G. C. & Lindahl, F. W., Stock Price Reaction to LIFO Adopters: The Association Between Excess Returns and LIFO Tax Savings, *Journal of Accounting Research* (Autumn 1982) pp. 551–588.

Brown, R. M., Short-range Market Reaction to Changes to LIFO Accounting Using Preliminary Earnings Announcement Dates, *Journal of Accounting Research* (Spring 1980) pp. 38–63.

Coleman, J. S., *Introduction to Mathematical Sociology* (New York: Free Press, 1964).

Comiskey, E. E. & Groves, R. E. V., The Adoption and Diffusion of an Accounting Innovation, *Accounting and Business Research* (Winter 1972) pp. 67–77.

Copeland, R. M. & Shank, J. K., LIFO and the Diffusion of Innovation, *Empirical Research in Accounting: Selected Studies* (1971) pp. 196–224.

Dirsmith, M. W., Discussion of Social Environments and Organizational Accounting, *Accounting, Organizations and Society* (1986) pp. 357–367.

DiMaggio, P. J. & Powell, W. W., The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields, *American Sociological Review* (April 1983) pp. 147–160.

Dodd, S. C., Formulas for Spreading Opinions, *Public Opinion Quarterly* (1958) pp. 537–554.

Fourt, L. A. & Woodlock, J. W., Early Prediction of Market Success for New Grocery Products, *Journal of Marketing* (1960) pp. 31–38.

Hagerman, R. L. & Zmijewski, M. E., Some Economic Determinants of Accounting Policy Choice, *Journal of Accounting and Economics* (1979) pp. 141–161.

Hicks, J. A., An Examination of Accounting Interest Groups' Differential Perceptions of Innovations, *The Accounting Review* (April 1978) pp. 371–388.

Hunt, H. G., Potential Determinants of Corporate Inventory Accounting Decisions, *Journal of Accounting Research* (Autumn 1985) pp. 448–467.

Hussein, M. E., The Innovative Process in Financial Accounting Standards Setting, *Accounting, Organizations and Society* (1981) pp. 27–37.

Kalwani, M. U. & Morrison, D. G., Estimating the Proportion of Always Buy and Never Buy Consumers: A Likelihood Ratio Test with Sample Size Implications, *Journal of Marketing Research* (1977) pp. 601–606.

Kennedy, P., *A Guide to Econometrics* (Cambridge: The MIT Press, 1979).

Leftwich, R., Accounting Information in Private Markets: Evidence from Private Lending Agreements, *The Accounting Review* (January 1983) pp. 23–42.

Lindahl, F. W., Accounting Change as a Dynamic Process, Working Paper, Duke University (January 1986).

Maddala, G. S., *Econometrics* (New York: McGraw-Hill, 1977).

Mahajan, V. & Peterson, R. A., *Models for Innovation Diffusion* (Beverly Hills: Sage Publications, 1985).

Mahajan, V. & Schoeman, M. E. F., Generalized Model for the Time Pattern of the Diffusion Process, *IEEE Transactions on Engineering Management* (February 1977) pp. 12–18.

Mansfield, E., *Industrial Research and Technological Innovation* (New York: Norton, 1968).

Mincer, J. & Zarnowitz, V., The Evaluation of Economic Forecasts, in Mincer, J. (ed.) *Economic Forecasts and Expectations*, pp. 3–46 (National Bureau of Economic Research, 1969).

Morse, D. & Richardson, G., The LIFO-FIFO Decision, *Journal of Accounting Research* (Spring 1983) pp. 106–127.

Ricks, W., The Market's Response to the 1974 LIFO Adoptions, *Journal of Accounting Research* (Autumn 1982) pp. 367–387.

Rogers, E. M., *Diffusion of Innovations* (New York: Free Press, 1962).

Rogers, E. M. & Shoemaker, F., *Communication of Innovation: A Cross Cultural Approach* (New York: Free Press, 1971).

Shank, J. K. & Copeland, R. M., Corporate Personality Theory and Changes in Accounting Methods: An Empirical Test, *The Accounting Review* (July 1973) pp. 494–501.

Sharif, M. N. & Ramanathan, K., Temporal Models of Innovation Diffusion, *IEEE Transactions on Engineering Management* (May 1984) pp. 76–86.

Sunder, S., Relationship Between Accounting Changes and Stock Prices: Problems of Measurement and Some Empirical Evidence, *Journal of Accounting Research* (1973) pp. 1–45.

Sunder, S., Stock Price and Risk Related to Accounting Changes in Inventory Valuation, *The Accounting Review* (April 1975) pp. 305–316.

Tritschler, C. A., A Sociological Perspective on Accounting Innovation, *The International Journal of Accounting Education and Research* (Spring 1970) pp. 39–67.

Watts, R. L. & Zimmerman, J. L., *Positive Accounting Theory* (Englewood Cliffs: Prentice-Hall, 1986).

Watts, R. L. & Zimmerman, J. L., Toward a Positive Theory of the Determination of Accounting Standards, *The Accounting Review* (January 1978) pp. 112–134.

Zmijewski, M. E. & Hagerman, R. L., An Income Strategy Approach to the Positive Theory of Accounting Standard Setting/Choice, *Journal of Accounting and Economics* (August 1981) pp. 129–149.

Accounting, Organizations and Society, Vol. 14, No. 4, pp. 321–324, 1989.
Printed in Great Britain

0361–3682/89 $3.00+.00
Pergamon Press plc

BUDGET EMPHASIS, BUDGETARY PARTICIPATION AND MANAGERIAL PERFORMANCE: A NOTE*

ALAN S. DUNK
School of Economic and Financial Studies, Macquarie University

Abstract

Ongoing research into the impact of budget emphasis on the relationship between participation and managerial performance has failed to provide consistent results. The purpose of this study was to conduct a replication of this two-way interaction model, drawing on data from a range of organizations in order to promote result generalizability and to test whether the model fits the data. It is proposed that a potential reason for why the results of this study conflict with those of prior research is that their findings were based on non-random sampling.

The recent paper by Brownell & Hirst (1986) has been the latest step in a series of studies addressing the question of whether or not superior managers' evaluation styles have an impact on the relationship between participation and subordinate manager performance. The sequence commenced with Hopwood (1972) speculating in his examination of cost center managers that evaluation styles focusing on accounting measures would have a negative effect on managerial performance. Otley (1978), in a study based on profit center managers could not corroborate Hopwood's supposition, since his results suggested that a leadership style emphasizing budget attainment was related to greater levels of performance. Brownell (1982), in an effort to reconcile the Hopwood (1972) and Otley (1978) results, found that a high budget emphasis is effective in terms of managerial performance under conditions of high budgetary participation but not under low participation. Brownell's results were based on responses from cost center managers drawn from one U.S. manufacturing company. Finally, Brownell & Hirst (1986) attempted to reconcile the conflicting results in a study predicated on cost and revenue center managers' responses drawn from one manufacturing organization in Australia. However, their results did not support the proposition of there being a significant three-way interaction between budget emphasis, budgetary participation and task uncertainty affecting managerial performance. Neither could the study confirm Brownell's earlier two-way interaction finding.

The purpose of this study was to attempt to resolve at least partially the conflict between the Brownell (1982) and the Brownell & Hirst (1986) results by conducting a replication of the Brownell two-way interaction model, examining the impact of budget emphasis on the relationship between participation and performance. Kerlinger & Pedhazur (1973, p. 447) emphasized the need for replication studies, since replication can enhance the external validity and generalizability of prior research if previous findings are subsequently supported.

In conducting the replication, two specific issues received close consideration. Firstly, Brownell & Hirst suggested that cross-national differences in samples (U.S. versus Australia) could explain why the performance hypotheses were not supported, but they could not develop an argument to explain why this might be so. There is no evidence in the literature that would infer that the cultural environment in which manufacturing firms operate in these countries is different in any way. Hofstede's (1984) results have shown that there is virtually no difference between samples from the U.S.A., Australia and the U.K. on his four dimensions of culture. Since cross-national differences were not expected to drive research results in these three countries, this study reports findings from British firms.

---

*Thanks are due to Peter Brownell, Graeme Harrison and Hwang Soo Chiat for their comments on earlier drafts of this paper.

Secondly, the research undertaken by Hopwood (1972), Otley (1978), Brownell (1982) and Brownell & Hirst (1986) all relied on data drawn from single organizations. As Brownell commented, the generalizability of his results may have been hampered by not only the non-random selection procedure within the subject organization, but also relying on data from only one firm. Since the validity of inferences made from a sample to a population is dependent upon the sample being representative of that population (e.g. Hopkins & Glass, 1978, p. 183), sample selection requires careful consideration. In order to ensure the reliability of research results, representative samples should be employed (Kerlinger & Pedhazur, 1973, p. 446). A randomly drawn sample is representative of the population (Hopkins & Glass, 1978, p. 185). In order to promote result generalizability and to limit the possibility of inadvertently controlling for some firm effect(s), a number of manufacturing organizations were sampled.

Since Brownell's model provided the basic framework for this research, the hypothesis underlying this study was as follows.

H_0: there is no significant interaction between budgetary participation and budget emphasis affecting performance.

METHOD

Production managers in each of 30 firms randomly selected from consumer product manufacturing organizations located in the north of Britain were personally requested to complete a mail questionnaire which asked for responses to three variables: budgetary participation, managerial performance and budget emphasis. Questionnaires were received from 26 managers, representing a response rate of 87%.

MEASURES

The Mahoney *et al.* (1963, 1965) manager self-rating measure was used to assess managerial performance. The measure has eight performance dimensions and one overall effectiveness dimension. It was found in this study that the eight dimensions explained 47% of the variance of the overall effectiveness dimension, a result consistent with Mahoney *et al.*'s findings. Data analysis was based on the overall effectiveness dimension, a procedure consistent with both Brownell (1982) and Brownell & Hirst (1986). Descriptive statistics for the measure are contained in Table 1.

Budgetary participation was assessed using the Milani (1975) six-item measure. Cronbach alpha (Cronbach, 1951) for the instrument was 0.867. Table 1 contains descriptive statistics for the measure.

To assess budget emphasis, Hopwood's (1972) eight-item supervisor evaluation style measure was employed. Evaluative style was determined by Hopwood by the presence or absence of the criteria of meeting the budget and concern with costs among the top three criteria as ranked by respondents out of the set of eight.

TABLE 1. Descriptive statistics for independent and dependent variables

Variable	Mean	Standard deviation	Theoretical range Min.	Max.	Actual range Min.	Max.
Performance	3.923	0.484	1	5	3	5
Budgetary participation	23.423	4.149	6	30	15	30
Budget emphasis	9.385	0.752	2	10	8	10

Since many of Brownell's (1982) respondents failed to complete the rank ordering properly, Brownell (1985) used, as was the case in this study, a five-point Likert scale for each of the eight criteria and the overall score for this measure was based on the sum of the two criteria of meeting the budget and concern with costs. Table 1 presents the descriptive statistics for this measure.

RESULTS

The following model was employed to test the hypothesis.

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + e$$

where Y is managerial performance, X_1 is budgetary participation (difference scores from mean, $\bar{P}-P$) and X_2 is budget emphasis: −1 for high budget emphasis and +1 for low budget emphasis (dichotomised at the mean).

Table 2 presents the results of the regression. Since the coefficient of the interaction term, b_3, is statistically significant, the null hypothesis of no interaction is rejected at $p < 0.05$. However, the rejection is not in the expected direction. The negative sign associated with b_3 suggests that high(low) participation together with high(low) budget emphasis lowers managerial performance, rather than increasing it.[1] Hence, the study confirms Brownell's result of a significant two-way interaction between participation and budget emphasis affecting performance, though it fails to support Brownell's finding that the interaction positively affects performance.[2]

CONCLUSION

In essence, results have now been produced that have shown a significant positive interaction (Brownell, 1982), no interaction (Brownell & Hirst, 1986) and now a significant negative interaction between participation and budget emphasis affecting performance. The finding of this study that performance is enhanced in conditions of high(low) participation and low(high) budget emphasis, indicates that alternate pairings of these variables should be avoided in the design of control systems, at least in the population sampled.

Recall that Brownell raised the point that the generalizability of his results may have been hampered by relying on data from only one firm. Sampling based on single versus multiple organizations could have been responsible for the variance in results between this study and those of Brownell's and Brownell & Hirst's. Based on the

TABLE 2. Regression results from testing the hypothesis

Variable	Coefficient	Value	Standard error	t	p
Intercept	b_0	3.923	0.091	43.14	<0.01
Budgetary participation	b_1	0.007	0.023	0.30	n.s.
Budget emphasis	b_2	−0.006	0.091	−0.07	n.s.
Participation × budget emphasis	b_3	−0.050	0.023	−2.19	<0.05

n.s. — not significant.
$R^2 = 19.6\%$, $n = 26$.

---

[1] The observed values for budget emphasis were restricted to 8, 9 and 10. However, random sampling provides some confidence that the range of scores in the population for this variable was captured. The dichotomisation treated scores of 8 and 9 as low and 10 as high. The alternative of treating 8 as low and 9 and 10 as high produces largely identical results, in that the interaction is still negative and significant ($t = -2.54$, $p < 0.02$).

[2] Given these differences in results, an examination was conducted to determine whether or not the linear model proposed to fit the data did not contravene the assumptions of regression analysis (Neter *et al*, 1985; Hirsch & Louderback, 1986). It was found that the specifications of the model as presented fit the data well in this study.

P, 6007

literature that argues for the need to employ representative samples (Kerlinger & Pedhazur, 1973; Hopkins & Glass, 1978), this may be the case. Nevertheless, further multiorganizational research is required to clarify whether or not this is the critical issue.

BIBLIOGRAPHY

Brownell, P., The Role of Accounting Data in Performance Evaluation, Budgetary Participation, and Organizational Effectiveness, *Journal of Accounting Research* (Spring 1982) pp. 12–27.

Brownell, P., Budgetary Systems and the Control of Functionally Differentiated Organizational Activites, *Journal of Accounting Research* (Autumn 1985) pp. 502–512.

Brownell, P. & Hirst, M., Reliance on Accounting Information, Budgetary Participation, and Task Uncertainty: Tests of a Three-way Interaction, *Journal of Accounting Research* (Autumn 1986) pp. 241–249.

Cronbach, L. J., Coefficient Alpha and the Internal Structure of Tests, *Psychometrika* (1951) pp. 297–334.

Hirsch, M. L. & Louderback III, J. G., *Cost Accounting: Accumulation, Analysis, and Use* (Boston: Kent Publishing Co., 1986).

Hofstede, G., *Culture's Consequences* (Beverly Hills: Sage Publications, 1984).

Hopkins, K. D. & Glass, G. V., *Basic Statistics for the Behavioral Sciences* (Englewood Cliffs: Prentice-Hall, 1978).

Hopwood, A. G., An Empirical Study of the Role of Accounting Data in Performance Evaluation, *Journal of Accounting Research* (Supplement 1972) pp. 156–182.

Kerlinger, F. N. & Pedhazur, E. J., *Multiple Regression in Behavioral Research* (New York: Holt, Rinehart and Winston, 1973).

Mahoney, T. A., Jerdee, T. H. & Carroll, S. J., *Development of Managerial Performance: A Research Approach* (Cincinnati: South-Western Publishing Co, 1963).

Mahoney, T. A., Jerdee, T. H. & Carroll, S. J., The Jobs of Management, *Industrial Relations* (February 1965) pp. 97–110.

Milani, K. W., The Relationship of Participation in Budget-setting to Industrial Supervisor Performance and Attitudes: A Field Study, *The Accounting Review* (April 1975) pp. 274–284.

Neter, J., Wasserman, W. & Kutner, M. H., *Applied Linear Statistical Models*, 2nd edn (Homewood, IL: Irwin, 1985).

Otley, D. T., Budget Use and Managerial Performance, *Journal of Accounting Research* (Spring 1978) pp. 122–149.

657.01
Ac 27
V.14

Accounting, Organizations and Society, Vol. 14, No. 4, pp. 325–335, 1989.
Printed in Great Britain

0361–3682/89 $3.00+.00
Pergamon Press plc

THE ROLE OF BUDGET DATA IN THE EVALUATION OF MANAGERIAL PERFORMANCE*

OLUMHENSE A. IMOISILI
Department of Accounting, University of Lagos

Abstract

This study was designed to test two of the contingency frameworks suggested by Otley (*Journal of Accounting Research*, 1978, pp. 122–149) as a possible explanation for the differences in results between his study and that of Hopwood (*Supplement to Journal of Accounting Research*, 1972, pp. 156–193). Using three manufacturing organizations in Northeastern U.S.A., it was found that the two contingency frameworks may not provide adequate explanation for the differences.

The role of budget data in the evaluation of managerial performance and the determination of rewards in organizations has been emphasized in the accounting literature. Researchers have formulated and/or tested hypotheses concerning the consequences of the different ways or "styles" that budget data may be used in performance evaluation and reward (Hopwood, 1972; Otley, 1978; Brownell, 1982). At best, the research results are inconsistent and inconclusive, however. For example, Hopwood (1972), using an American firm, found that managers evaluated on a style which used budget data in a rigid manner in performance evaluation reported significantly higher levels of job tension than managers evaluated on their contribution to organizational effectiveness. Otley (1978) replicated Hopwood's study in a British firm. He found no significant differences in the levels of job tension and performance reported by managers evaluated on the styles initially used by Hopwood. Faced with such differences, other researchers have formulated contingency frameworks where the effectiveness of using budget data in performance evaluation and reward is hypothesized to depend on the context in which the styles are used (Hayes, 1977; Otley, 1980; Birnberg *et al.*, 1983).

Otley (1978, 1980) suggested a style–context framework to reconcile the differences in results between his study and that of Hopwood (1972). The style dimension consisted of a high and a low emphasis on budget data in performance evaluation. The context dimension consisted of managerial interdependency[1] (high and low) and task uncertainty[2] (also high and low). Otley's framework suggested that when managers face high levels of interdependency or uncertainty, a high emphasis on budget data in performance evaluation may be dysfunctional for

---

*The author wishes to extend his appreciation to his dissertation committee members, Jake Birnberg (Chairman), Barry Lewis, Craig Russell, Larry Feick and Evelyn Perloff for their support throughout the duration of this research. The author's appreciation also goes to Neil Wilner, J. Harry Evans, III, Jeff Pinto, Stan Biggs, Ken Merchant and two anonymous referees for their comments and suggestions on an earlier draft of this paper.

[1] Managerial interdependency is the extent to which each manager perceives his or her work-related activities to require the joint or cooperative effort of other managers within that organization (Osborne *et al.*, 1980).

[2] Task uncertainty is the extent to which managers can predict confidently the factors which have effect on their work-related activities (Duncan, 1972; Otley, 1978).

managerial performance. A high emphasis on budget is more effective than a low emphasis in contexts characterized by low levels of task interdependency or uncertainty. However, although suggestive, such a contingency formulation has not been subjected to rigorous empirical analysis.

The current research was designed to test the two contingency models formulated by Otley (1980). The study tested the effectiveness of using budget data in performance evaluation and reward in task contexts characterized by high or low levels of interdependency or uncertainty.

Based on the above frameworks, the rest of the paper is organized in the following manner: the next section provides a review of the literature and a specification of the hypotheses tested. The research method and results are presented thereafter. This is followed by discussion and conclusion.

LITERATURE REVIEW AND HYPOTHESES

Previous research in accounting has identified some of the consequences of using budget data in performance evaluation. The arguments and findings suggest that when a high emphasis is put on using budget data for performance evaluation and reward, managers may report job tension and behave in ways which are dysfunctional for organizational goals. However, the arguments and conclusions were based on inconsistent and inconclusive empirical-results or case evidence (Argyris, 1952; Simon *et al.*, 1954; Hofstede, 1967, Lawler, 1976).

Hopwood (1972) was the first researcher to systematically test for the consequences of the different styles of using budget data in performance evaluation. The three styles he used in his empirical test were: (1) a budget constrained style (BCS) where budget data was used in a rigid manner in performance evaluation; (2) a profit conscious style (PCS) where the emphasis in performance evaluation was on contribution to organizational profitability; and (3) a non-accounting style (NAS) where budget data played a relatively unimportant role in performance evaluation. Hopwood tested for the consequences of these styles and found the BCS, compared to PCS or NAS led to greater stress, more hostility toward supervisors and peers, a greater tendency to blame other managers for poor performance outcomes and an increased tendency to charge departmental costs to other departments. He also reported that BCS managers were more likely than PCS or NAS ones to falsify their budget data and were less likely to meet their budgets. Based on the differential effect of the three styles on stress, Hopwood suggested that BCS would have a more negative impact on managerial performance than PCS or NAS.

Otley (1978) replicated Hopwood's study in a British firm that was well suited to the use of PCS in performance evaluation. He did not confirm Hopwood's initial results. There were no significant differences in the levels of performance and stress reported by managers evaluated on BCS and those evaluated on PCS. Otley's BCS managers were more likely to meet their budgets than his PCS managers, a result that was opposite to that of Hopwood. Three explanations were offered for the differences in results. First, Otley's managers were said to operate more independently of other units within the same organization than Hopwood's managers. Secondly, Otley's managers were profit center managers whereas Hopwood's subjects were cost center managers. Finally, Hopwood's managers were said to operate in a less predictable environment than Otley's managers.

The preceding review suggests some critical questions to ask: why did Hopwood's BCS managers report more stress than his PCS managers whereas Otley found no significant differences in the level of stress reported by his BCS and PCS managers? If one style of evaluating performance leads to more stress, does that style also lead to lower levels of managerial performance? Under what conditions do the different performance evaluation styles lead the managers to report favorable attitudes towards their budgets?

These questions are interesting to ask because organizations may show an interest in the differences in the performance evaluation styles if it can be consistently demonstrated that one style enhances performance more than another. Secondly, the effects of the different styles of evaluating managerial performance on attitude towards the budget may be interesting to study since it has been suggested that managerial attitudes towards the budget may influence how managers use budget data in decision-making and problem-solving (Birnberg *et al.*, 1983). The more favorable the attitude, the more confidence managers may have in using budget data in decision-making.

The effects of the interaction between budget evaluation style and task interdependency or uncertainty on stress, performance and managerial attitude have been given limited attention in empirical research in accounting. Only one study has investigated the appropriateness of using accounting data in performance evaluation in task contexts characterized by different levels of interdependency (Hayes, 1977). In that study, interdependency was not measured empirically but was assumed. Similarly, few research studies have addressed the effectiveness of accounting data in performance evaluation when task uncertainty is controlled (Rockness & Shields, 1984; Gordon & Narayanan, 1984). In all of these studies, no consistent result has emerged.

It is argued in this study that because tasks involving interdependency or uncertainty may be subject to evaluation using budget data, managers may be concerned with how such budget data are used for their evaluation and reward. In tasks characterized by high uncertainty or interdependency, managers may perceive themselves as having less than full control over performance outcomes. Using budget data in a rigid manner in such situations may be dysfunctional for performance since a rigid use of budget data assumes that most of the factors which have an effect on task outcomes are within the control of the managers being evaluated (Ansari, 1979). On the other hand, a rigid use of budget data may be more acceptable to managers if they perceive themselves able to exercise control over their performance outcomes. This would be the case in tasks characterized by low uncertainty or interdependency. In other words, it is not budget style *per se* which may lead to higher stress or lower performance. Rather, it is the mismatch of budget style and task contexts which may enhance stress or reduce managerial performance. The arguments presented above formed the basis for the following hypothesis:

> There will be significant interactions between budget style of performance evaluation and task interdependency on the one hand and between budget style and task uncertainty on the other, on job stress, performance and attitude toward the budget, i.e. when tasks are characterized as high on interdependency or uncertainty a BCS used in performance evaluation will lead to greater stress, lower performance and less favorable attitude toward the budget than the use of a PCS. On the other hand, when tasks are low on interdependency or uncertainty, a BCS used in performance evaluation will lead to higher performance, lower stress and a more favorable attitude toward the budget than the use of a PCS.

RESEARCH METHOD

The research method was a field study using a questionnaire survey. Three organizations participated in the study. The organizations included a tool manufacturing company with a sales volume of $460 million and an employee size of about 4000 in 1983, an equipment division of a manufacturing company with sales of $1 billion and an employee size of 9700 in 1983 and a hospital manufacturing company with sales of $300 million and an employee size of 3400 in 1983.[3]

Participants in the study were cost center managers[4] with budget responsibilities ranging from $200,000 to $1.5 million a year. The participants had been in their job positions for an average of 5 years. Each had also worked for the organization they were employed in for an average of 13 years.

---

[3] The first two companies had reduced their employee size by almost 40% between 1982 and 1983. Both had reported financial losses for 8 consecutive quarters prior to the time the research was started. The third company had reported declining profits for 4 consecutive quarters prior to the time of the study.

[4] Although attempts were made to include profit center managers in the study, two of the participating organizations would only allow their cost center managers to participate in the study. Since the third company did not have enough profit center managers to act as a control sample, it was decided to restrict the study to cost center managers. The bias this choice of managers could possibly have on the results of this study should be noted when interpreting the results.

Research procedures

The participants were identified by the researcher and the coordinator in each organization using the organization chart.[5] The criteria used in selecting the participants included: (1) each participant should have budget responsibility; (2) each person must have been in that particular position for at least 2 years; (3) two participants with the same supervisor should not participate in the study simultaneously; and (4) no manager who had participated in the pilot test phase should participate in completing the questionnaire. The research was undertaken in three phases.

Phase 1. An unstructured interview was undertaken in two[6] of the participating organizations. Managers with budget responsibilities were used for these interviews. The purpose of this phase was 4-fold: (1) to gain an insight into the organization's budgeting process; (2) to determine the appropriate level of management to participate in the study; (3) to identify how the budget styles used in the research manifested themselves in the organizations which participated in the study; and (4) to identify whether there were other important issues or problems faced by the participants besides the ones studied.

Phase 2. Some of the measures used in the study were pretested in two of the three organizations which participated. The pretest subjects were encouraged to put comments on any items which were not clear to them. Based on the pretest, some items were reworded, others were eliminated, while a few were added. In general, the interpretation of most of the items by the participants was consistent with the way the researcher wanted them interpreted.

Phase 3. One hundred packages were sent to each of the three organizations to distribute to participating managers. Each package contained two questionnaires (one for the participant and the second for the supervisor of the participant), two prepaid postal envelopes to be used to return the questionnaires directly to the researcher and two letters. One letter was from the researcher explaining the purpose of the research and soliciting their participation. The second letter was from a Vice-President of the participating company explaining that the questionnaires would be used for research purposes only. They were informed in the Vice-President's letter than participation was voluntary. The questionnaire itself contained statements assuring the participants that their responses would be treated with absolute confidentiality. Each package was sent to the supervisor or boss who was expected to complete his or her part and give the second part to the subordinate so named by the Vice-President in his memo.

Table 1 contains the responses of managers from each participating organization. Although 300 questionnaires were given to the coordinators to distribute among their managerial employees deemed appropriate for the study, a total of 188 questionnaires were actually given out in the three organizations.[7]

---

[5] The coordinators were appointed by each company on the basis of their familiarity with the budgeting process in that particular organization.

[6] The third company would not agree for an extensive interview to be conducted. However, interviews were held with some of the budgeting staff and two Vice-Presidents to confirm that budget was used in that company in a manner deemed appropriate for this study. The company was included in the study when it was found that their budgeting procedures were consistent with prior expectation.

[7] The coordinators only gave the questionnaires to those managers and their supervisors who agreed to participate in the study when they were contacted on the telephone. Since the coordinators did not keep any record of those managers who refused participation, it was not possible to find out how much non-participation would bias the results of the study.

TABLE 1. Responses by organization

	Organization 1	Organization 2	Organization 3	Total
Total questionnaire*	100	100	100	300
Total given out	45	76	67	188
Number of responses	27	52	41	120
Response rate†	60%	68%	61%	64%
Number usable	24	41	37	102
Usable response rate†	53%	54%	58%	54%

*Number delivered to the coordinators in each organization.
†As a percentage of questionnaires given out.

Of the number distributed, 120 were returned[8] giving a response rate of 64%. Eighteen questionnaires were dropped from the analysis either because all parts of the questionnaires were not completed or because the respondents confirmed in one statement that they did not have budget responsibilities. One hundred and two questionnaires representing 54% of the total given out were used for analysis.

Measurement and validation of variables

Task interdependency. Task interdependency was defined as the extent to which each participant perceives his or her work-related activities to require the joint or cooperative efforts of other work-groups within that particular organization. A 2-item, 5-point, Likert scale from Osborne *et al.* (1980) was used in the study.

Task uncertainty. Task uncertainty was defined as frequency of change of factors affecting work-unit performance (Duncan, 1972). A 5-item, 5-point, Likert scale from Podsakoff *et al.* (1983) was used to measure the variable.

Job stress. Stress was defined as the extent to which an individual faces situations characterized by incompatible demands on his or her role performance or when two or more individuals make incompatible or inconsistent demands on the focal individual (Rizzo *et al.*, 1970). A 7-item, 7-point, Likert scale from Rizzo *et al.* (1970) was used for the study. The instrument used for this research was different from that used by Hopwood (1972) and Otley (1978).

Performance. Seven dimensions of managerial performance used for empirical analysis by Porter & Lawler (1968) and Van de Ven & Ferry (1980) were used for this research. The dimensions include quality of work, attitude towards work, effort, speed of work, frequency of useful ideas suggested, frequency of meeting work goals and cost efficiency of the work unit of which the particular manager is head. These dimensions were chosen because they were included in the performance appraisal forms used by the three companies for evaluating the performance of managerial employees. The supervisors of the participating managers were required to rate the participants on a 7-point Likert scale ranging from "well below average" to "well above average" on the 7 dimensions taking into consideration all subordinates he or she had supervised over the previous 3 years.

Attitude towards the budget. Respondents were required to indicate how favourable or unfavourable their attitudes were towards the budgeting staffs. The scale was measured using a 2-item, 7-point, Likert scale made for the study.

---

[8] Sixty-eight manager/supervisor pairs who had initially agreed to participate in the study did not return their questionnaires. This number includes managers in two departments where the Vice-Presidents who had given prior approval to their departmental participation in the study later changed their minds. These two VPs could not be contacted to know why they vetoed their departmental participation. The other managers who did not return their questionnaires likewise could not be contacted because the coordinators were precluded by their organizations from allowing the researcher to deal directly with participating managers.

Budget evaluation styles. The styles initially used by Hopwood (1972) and Otley (1978) were used for this study. They included (1) budget constrained style (BCS), (2) profit conscious style (PCS)[9] and (3) non-accounting style (NAS).[10]

Eight items developed by Hopwood (1972) were used[11] for data analysis. Participating managers were required to indicate on a 7-point Likert scale the extent to which they perceived their supervisors to use the 8 items in their performance evaluation. Each manager was assigned to the style on which he or she had the highest average score. Fifty-eight managers were classified as being evaluated on BCS,[12] 40 on PCS and 4 on NAS. The 4 managers who reported that they were evaluated on NAS were dropped from the analysis because this group were not sufficiently numerous for the empirical tests.

The descriptive statistics and internal reliability of the variables used in the analysis are summarized in Table 2. Cronbach coefficient alpha (Cronbach, 1951) was used to determine internal reliability of each multi-item variable.

RESULTS

The following equation was tested using multi-variate linear regression (SPSSx, 1986, pp. 527–533):

$$y_j = b_{0j} + b_{1j}Z + b_{2j}X_1 + b_{1j}X_2 + b_{4j}X_1Z + b_{5j}X_2Z + b_{6j}X_1X_2 + b_{7j}X_1X_2Z + e_{ij}$$

where

$j = 1, 2, 3$,

Y_1 = job stress,

TABLE 2. Descriptive statistics of variables used in empirical analysis

	No. of Items*	Theoretical range	Actual range	Mean	Standard deviation	Cronbach's alpha
1. Task interdependency	2	2–10	2–10	7.3	1.8	0.77
2. Task uncertainty	5	5–25	9–25	16.0	3.3	0.79
3. Job stress	7	7–49	14–47	28.1	6.3	0.73
4. Performance	7	7–49	22–42	32.7	5.0	0.94
5. Attitude towards the budget	2	2–14	2–14	9.0	2.4	0 86

*No. of items used for empirical analysis.

---

[9] In the two organizations in which the interviews were carried out, PCS manifested itself in the form of how cost effective the respective managers were in their operations. Hopwood (1972) also reported this concern in his study.

[10] Although budget data may not be considered as important in performance evaluation, the style does not preclude the use of budget information for other purposes such as problem-solving. This study only focused attention on one type of use to which budget data may be put, i.e. in evaluating performance and determining rewards in organizations.

[11] These 8 items were included in a list of 21 items given to the managers and their supervisors to respond to. The 21 items also included 13 others generated during the interview with some of the participating managers. All of the 21 items combined represent what the managers considered as most important in their performance evaluation. The list was also checked with some supervisors and human resources managers to ensure that the items were truly representative of the criteria used by the companies in performance evaluation.

[12] Both Hopwood (1972) and Otley (1978) asked participating managers in their studies to rank-order the items they used in measuring their budget styles. The ranks were used to derive the three styles they used in empirical analysis. This research used a simple average of the raw scores to determine the styles used in empirical analysis. The differences in the way style was measured between this study, on the one hand, and Hopwood (1972) and Otley (1978) on the other, should be noted in interpreting the results reported in this study. It is possible that differences in measurement may have influenced the results reported in this study.

Y_2 = performance,
Y_3 = attitude towards the budget,
X_1 = task interdependency,
X_2 = task uncertainty, and
Z = budget style (code 1 for BSC, 0 for PCS).

Multi-variate linear regression is appropriate to use when there are two or more dependent variables being regressed on one or more independent variables simultaneously. In the test carried out, performance, job stress and attitude towards the budget were used as dependent variables. The independent variables included budget style (coded 1 for BCS and 0 for PCS), task interdependency, task uncertainty as well as interaction terms between budget style and task interdependency, budget style and task uncertainty, task uncertainty and task interdependency and budget style, task interdependency and task uncertainty.

The results of the test of the hypotheses are reported in Table 3. The regression coefficients representing a test of hypotheses are presented by β_{4j} and β_{5j} in Table 3. As can be observed from the table, none of the six regression coefficients tested was significant at the 5% alpha level. This means that none of the interaction terms of budget style and task uncertainty as well as budget style and task interdependency on performance, job stress and attitude towards the budget was significant. Managers who reported that they were evaluated on BCS and who operated in task contexts characterized as high or low on interdependency or uncertainty did not report higher performance, job stress or more favorable attitudes toward the budget than PCS managers.

Also from Table 3, it can be seen that the regression coefficients from budget style to job stress and from budget style to attitude towards the budget were significant at an alpha level of 5%. To test for the direction of significance, a t-test was carried out with job stress and attitude towards the budget as the dependent variables and budget style (coded 1 for BCS and 0 for PCS) as the independent variable. It was found that BCS managers reported higher job stress than PCS managers. This result is consistent with that reported by Hopwood (1972) as well as the related conceptual literature (Hirst, 1981). BCS managers also reported a more favorable attitude towards the budget than PCS managers. This finding is similar to that of Otley (1978) who reported that BCS managers were more likely to meet their budgets than PCS managers.

DISCUSSION

The study focused on the effects of budget

TABLE 3. MRLA test of interaction effect of budget style and task interdependency or uncertainty on job stress, performance and attitude towards the budget

	Dependent variables		
	Job stress	Performance	Attitude towards the budget
Direct effects of:			
Budget style (BS): β_{1j}	1.67†	ns	−1.24†
Task interdependency (TI): β_{2j}	ns	ns	ns
Task uncertainty (TU): β_{3j}	ns	ns	ns
Interaction effects:			
BS × TI: β_{4j}	ns	ns	ns
BS * TU: β_{5j}	ns	ns	ns
TI * TU: β_{6j}	ns	ns	ns
BS × TI × TU: β_{7j}	ns	ns	ns
Uni-variate F	5.22§	ns	1.85*

*$p \leq 0.10$; †$p \leq 0.05$; ‡$p \leq 0.01$; §$p \leq 0.001$.
ns: not significant.

style of evaluating managerial performance on job stress, attitude towards the budget and performance in task contexts characterized as high or low on interdependency or uncertainty. As reported in the previous section, the hypotheses tested were not supported. In other words, the contingency frameworks suggested by Otley (1978) may not provide adequate explanation for the differences in results between Hopwood's study and that of Otley.

There are two other possible explanations for the differences in results between the two studies, however. These are the financial condition of the organization and reward uncertainty.

Financial condition of organization

As explained earlier, all three organizations which provided managers for this study had experienced financial losses or declining profitability for 4–8 consecutive quarters prior to the time this research was conducted. Perhaps when organizations are facing financial hardship, the differential effects of performance evaluation styles on work outcomes may be difficult to isolate empirically. It is likely that when companies are undergoing periods of declining profitability, there might be so much emphasis on financial data in performance evaluation that it may be difficult to truly classify managers into different styles on the basis of how much consideration is given to budget data in performance evaluation.[13] It is possible that the styles used for this research may only be found in organizations which have no severe pressures for financial performance.

That financial difficulties may explain failure to provide support for the hypotheses tested in this research is further illustrated by Table 4 which provides a comparison of the distribution of NAS in this study with those in Hopwood (1972), Otley (1978) and Brownell (1982). In the studies of Hopwood and Brownell where the hypotheses tested were supported, 44% and 31%, respectively, of the managers who participated in the studies were classified as being evaluated on NAS. On the other hand, in this current study and that of Otley where the hypotheses tested were not supported, the percentages of managers who reported they were evaluated on NAS were 4% and 3%, respectively.[14]

Perhaps when organizations are passing through financially difficult times, there are no managers with budget responsibilities who are not evaluated on the basis of budget data. There

TABLE 4. Comparison of NAS in this study with Hopwood, Otley and Brownell

	n	Based styles*	NAS	Hypotheses
Hopwood (1972)	167	56%	44%	Supported
Otley (1978)	39	97%	3%	Not supported
Brownell (1982)	38	69%	31%	Supported
This study	102	96%	4%	Not supported

*Budget based styles include Hopwood's BCS and PCS.

---

[13] Otley (1978) made a similar argument that the style used in performance evaluation may be due in part to economic circumstances. According to Otley, a profit center experiencing financial difficulty was more likely to formulate optimistic budget standards than another profit center experiencing reasonable profitability. This is because when a profit center is reporting declining profitability, the manager was more likely to be evaluated either on multiple criteria or in a flexible manner than when he or she was experiencing profitability. The explanation offered by this research that the financial condition of the organization as a whole may explain differences in results between the studies of Hopwood and Otley in fact supports Otley's suggestion that the economic situation of a profit center may determine how budget data are used in performance evaluation.

[14] Professor David Otley, in a private communication with this researcher confirmed that the organization he used in his study was actually experiencing financial difficulty at the time he conducted his research.

is also the possibility that during periods of poor financial performance, there might be so much emphasis on budget information and reports on the interactions between a manager and his or her supervisor that most managers may have a tendency to think they are evaluated on the basis of budget data even though that might not be the case in circumstances where the company was profitable.

That the financial status of the organizations used for analysis may have influenced the results of this study is further supported by the following quotation from an interview conducted in phase 2 of the research:

> You are told to prepare your budget... they tell you to be daring, innovative and creative... I spend weeks preparing my plans. When they send you your approved budget, they cut under your feet.

Another R & D manager also complained that:

> For example, based on your technical plan,[15] you budget for an equipment of $50,000. When it gets there,[16] they order 20% across the board cut. Meanwhile, you cannot buy 80% of an equipment. Either you buy it or you don't. Eventually, we end up with unworkable budget and technical plans.

It is likely that organizations which face financial pressures run their budget processes as described in the statements quoted above. It is probable that across the board cuts in all areas of operations are found in organizations experiencing financial difficulties.

Reward uncertainty

A second reason the hypotheses tested in this study have not been supported may be a result of the managers used for analysis not being aware of how they were evaluated by their supervisors. In this study, 21 performance evaluation criteria[17] were given to both managers and their supervisors to complete. The managers were required to indicate to what extent they thought their supervisors used the 21 items in their performance evaluation. The supervisors were also required to indicate the extent to which they actually used the 21 items in evaluating the manager jointly participating in the research with them. A Pearson correlation was run between both sets of responses. The correlations between both sets of responses for the overall sample, the BCS and PCS subsamples are reported in Table 5. Overall, 8 items were positive and significant, 3 were negative and significant and 11 were not significant. The comparative figures for BCS and PCS managers are summarized in the table.[18]

When budget style was defined on the basis of the responses of the superiors, the superiors reported that they put less emphasis on budget data in performance evaluation than their subordinates reported. For example, 26 out of 102 superiors reported that they evaluated their subordinates on BCS, while 31 and 45 reported PCS and NAS, respectively. The corresponding figures for the subordinates were 58, 40 and 4 for BCS, PCS and NAS, respectively.

That reward uncertainty may also be a possible explanation for failure to support the hypotheses tested in this research is further supported by the following statement made by one of the managers who participated in an interview phase of this research:[19]

---

[15] These are the research and product development plans and objectives of R & D managers.

[16] This refers to the top management levels where the final budget decisions are made.

[17] These 21 items include 8 items from Hopwood (1972). The additional items were generated from the interview stage.

[18] Hopwood (1972) also argued that the style a manager perceives himself to be evaluated on may be different from that actually used by his supervisor in performance evaluation. The result reported here is, in fact, support for Hopwood's argument that there may be disagreement between a manager and his supervisor on the criteria used in performance evaluation.

[19] The statement quoted here is typical of the complaints many of the managers who participated in the interview phase made.

TABLE 5. Distribution of correlation coefficients between the responses of managers and their supervisors on criteria used in evaluating managerial performance

	Overall sample	BCS	PCS
Significant and positive correlation coefficients	8	6	5
Significant and negative correlation coefficients	1	4	2
Not significant correlation coefficient	12	11	14
Total	21	21	21

> Information in this organization is top down ... You do not receive information as a matter of routine here. You receive them because you fight for them ... they only provide you with information when they want to beat up on you.

Many of the managers complained that their organizations did not have formalized processes for informing them what they had to do for them to be successful performers and, hence, rewarded accordingly.

CONCLUSION

This research sought to test the two contingency frameworks suggested by Otley (1978). The hypotheses tested were not supported suggesting that Otley's frameworks may not adequately explain the differences in results between his study and that of Hopwood.

Two reasons were offered as possible explanations for failure to support the hypotheses tested. One was the financial condition of the organizations used in the analysis. The three organizations which provided managerial employees for this study had experienced 4–8 quarters of financial losses or declining profitability prior to the time this current research was conducted. Secondly, the managers used in analysis may not have been aware of how their performances were evaluated and rewarded.

BIBLIOGRAPHY

Ansari, S., Towards an Open System Approach to Budgeting, *Accounting, Organizations and Society* (1979) pp. 149–161.

Argyris, C. *The Impact of Budgets on People* (New York: Controllership Foundation, 1952).

Birnberg, J., Turopolec, L. & Young, S. M., The Organizational Context of Accounting, *Accounting, Organizations and Society* (1983) pp. 111–129.

Brownell, P., The Role of Accounting Data in Performance Evaluation, Budgetary Participation and Organizational Effectiveness, *Journal of Accounting Research* (1982) pp. 12–37.

Cronbach, L., Coefficient Alpha and the Internal Structure of Tests, *Psychometrica* (1951) pp. 297–334.

Duncan, R., Characteristics of Organizational Environment and Perceived Environmental Uncertainty, *Administrative Science Quarterly* (1972) pp. 313–327.

Gordon, L. & Narayanan, V., Management Accounting Systems, Perceived Environmental Uncertainty and Organizational Structure: An Empirical Investigation, *Accounting, Organizations and Society* (1984) pp. 33–47.

Hayes, D., The Contingency Theory of Management Accounting, *The Accounting Review* (1977) pp. 22–39.

Hirst, M., Accounting Information and the Evaluation of Subordinate Performance, *The Accounting Review* (1981) pp. 771–784.

Hofstede, G., *The Game of Budget Control* (Assen: Van Gorcum Co., 1967).

Hopwood, A., An Empirical Study of the Role of Accounting Data in Performance Evaluation, *Supplement to Journal of Accounting Research* (1972) pp. 156–193.

Lawler, E., Control Systems in Organizations, in Dunnette, M. (ed.) *Handbook of Industrial and Organizational Psychology* (Chicago, IL: Rand McNally, 1976).

Osborne, R., Hunt, J. & Hunt, L., *Organization Theory: An Integrated Approach* (New York: Wiley & Sons, 1980).

Otley, D., Budget Use and Managerial Performance, *Journal of Accounting Research* (1978) pp. 122–149.

Otley, D., The Contingency of Theory of Management Accounting: Achievement and Prognosis, *Accounting, Organizations and Society* (1980) pp. 413–428.

Podsakoff, P., Todor, W. & Schuler, R., Leader Expertise as a Moderator of the Effects of Instrumental and Supportive Leader Behavior, *Journal of Management* (1983) pp. 173–186.

Porter, L. & Lawler, E., *Managerial Attitudes and Performance* (Homewood, IL: Irwin, 1968).

Rizzo, J., House, R. & Lirtzman, S., Role Conflict and Ambiguity in Complex Organizations, *Administrative Science Quarterly* (1970) pp. 150–163.

Rockness, H. & Shields, M., Organizational Control Systems in Research and Development, *Accounting, Organizations and Society* (1984) pp. 165–177.

Simon, H., Guetzkow, H., Kozmetsky, R. & Tyndall, T., *Centralization Versus Decentralization in Organizing The Controller's Department* (New York: Controllership Foundation, 1954).

SPSSx, *User's Guide* (New York: McGraw-Hill Book Co, 1986).

Van de Ven, A. & Ferry, D., *Measuring and Assessing Organizations* (New York: Wiley & Sons, 1980).

Accounting, Organizations and Society, Vol. 14, No. 4, pp. 337–345, 1989.
Printed in Great Britain

0361–3682/89 $3.00+.00
Pergamon Press plc

CONSENSUS AND ACCURACY IN ACCOUNTING STUDIES OF DECISION-MAKING: A NOTE ON A NEW MEASURE OF CONSENSUS

KEVIN KEASEY
School of Industrial and Business Studies, Warwick University

and

ROBERT WATSON
Manchester School of Management, UMIST

Abstract

Despite the common use of consensus amongst experts as a "surrogate for truth" in experimental studies of decision-making, there has been little discussion as to its appropriateness. A new measure of consensus for dichotomous predictions, which takes into account experts' confidence in the correctness of their decisions, is proposed and its relationship with accuracy is experimentally evaluated. The experiment also extends Ashton's (*The Accounting Review*, April 1985, pp. 173–185) empirical work on the conventional consensus measure by checking the robustness of Ashton's results to a different experimental setting (U.K. trade credit specialists) and to increases in the cues made available to experimental participants. The paper concludes that the new measure of consensus complements the conventional measure, in that it provides information over and above that provided by the conventional consensus measure. The paper also concludes that the results noted by Ashton for the U.S. also hold for the U.K. and are robust to increases in the cue set made available to participants.

Much of the experimental research into various aspects of individual decision-making has had a practical aim, namely, the making of "better" decisions. Irrespective of how "better" is defined in any particular context, some form of "benchmark" or criterion to determine whether or not one decision is superior to another is required. Ideally, some "objective" and measurable external criterion such as accuracy or "correctness" should be used. However, in situations where no objective external criteria exist, as in many evaluative and assessment type accounting and auditing decision situations, this is obviously infeasible. In these situations, recourse is usually made to the level of consensus between experts, where consensus is defined as the number of agreements between experts.

The purpose of this paper is to consider the appropriateness of consensus, as defined above, given the role it is assigned in experimental studies. After considering the role of consensus in experimental studies, a new measure of consensus is proposed. Both the conventional and new consensus measures are then experimentally evaluated using Ashton's (1985) procedure of comparing their relative correlations with accuracy. The paper also extends Ashton's (1985) empirical evaluation of the conventional consensus measure in two ways: first, by determining the robustness of her results in a different experimental setting, and second, by assessing the impact of increases in the cue set made available to experimental participants.

The paper contains the following sections. First, the role of consensus in experimental situations is considered and a new measure proposed. Second, the nature of the experiment used for the evaluation of the new measure is briefly described. The third section presents the empirical findings for both pairwise and individual definitions of consensus and accuracy. A discussion of the results and conclusions are given in the final section.

CONSENSUS IN EXPERIMENTAL STUDIES

Measures of consensus

The ultimate purpose of much experimental work on decision-making in accounting is to understand and improve decision-making in practice. As a consequence, the correctness of decisions made in experiments needs to be evaluated. In some decision-making situations there is an objective measure, such as accuracy, against which to compare the decisions made by participants in an experiment. For example, in experimental studies concerned with failure prediction, researchers are able to judge the correctness of decisions by their knowledge of which companies actually failed. However, a number of decision situations in accounting are evaluative in nature (for example, internal control evaluation) and, therefore, recourse has to be made to a different measure of the quality of decisions. In these types of situations the level of consensus between experts is the usual measure of decision quality,

> "Because we lack an objective definition of a 'correct' decision in situations such as materiality judgement, consensus judgements of experts are often employed as a substitute criterion .... In other situations such as internal control evaluation, the degree of judgemental consensus is often used as a substitute measure of decision quality; lack of consensus indicating that at least some individual's judgements are incorrect" (Libby & Lewis, 1982, p. 234).

However, this conventional measure of consensus may not provide a good proxy for accuracy in experimental situations, particularly where dichotomous predictions are required. This is because consensus merely records the amount of agreement across experts and does not differentiate between agreement over correct decisions and agreement over incorrect decisions. Thus, high levels of consensus need not imply "better" or more "correct" decisions since all subjects could be equally wrong. Whether or not consensus is an adequate surrogate for truth is, therefore, an entirely empirical matter.

Moreover, the conventional consensus measure appears to incorporate an excessively restrictive view of expertise. Arguably, a richer view of expertise would be where experts not only make correct decisions but are also reasonably able to judge *ex-ante* the correctness of their decisions. A consensus measure that includes experts' confidences in their decisions may, therefore, provide an improved proxy for accuracy. Before considering whether this is the case empirically, consideration will be given to the circumstances where confidence and accuracy might be well matched.

First, experimental results (for example, Clarke, 1960; Nickerson & McGoldrick, 1965; Pitz, 1974; Lichtenstein & Fischhoff, 1977, 1980a) suggest that the more difficult the experimental task, the more likely a subject's confidence is to outweigh his/her accuracy. In other words the general finding in experimental studies that subjects tend to be over-confident (for example, Oskamp, 1965; Hazard & Peterson, 1973; Koriat *et al.*, 1980) weakens as the task presented to subjects becomes easier. This "difficulty" effect has been explained by Lichtenstein *et al.* (1982) as arising from subjects' inability to assess the relative difficulty of the task facing them.

A second factor affecting the relationship between confidence and accuracy is the amount of feedback given to the subjects regarding their performance. Research by Adams & Adams (1958, 1961) and Lichtenstein & Fischhoff (1980a) found that extensive feedback concerning subjects' prior performance improved the subjects' subsequent performance. In contrast, Choo (1976) found that a single feedback session did little to improve performance. Einhorn (1982) argues, however,

> "outcome information, without knowledge of structure, can be irrelevant for providing self-correcting feedback about poor heuristics. It has also been argued that knowledge of task structure is difficult to achieve because of the inductive way in which we learn from experience" (p. 282).

How feedback affects the relationship between confidence and accuracy is, therefore, dependent upon the amount of feedback provided to subjects and the subjects' knowledge of the structure of the task facing them. In connection with this last point, a study by Koriat *et al.* (1980) noted that a task structure that promoted reflection on the decision to be made, significantly reduced over-confidence.

Third, it is to be expected that the motivation of subjects would affect how their confidence in a decision related to their accuracy. In a number of situations, for example, standard setting, the motivation of subjects could clearly lead to under-confident assessments. However, the limited existing research on this topic (Sieber, 1974) indicates that increased motivation does not lead to confidence and accuracy being better matched.

A fourth factor to be considered is the expertise of the subjects given the task structure at hand. Lichtenstein & Fischhoff (1977) found that their student subjects were no better "calibrated" (high calibration existing when confidence and accuracy are well-matched and vice versa) when answering questions in their field of study as compared to general knowledge questions. This lack of an expertise effect on calibration is mirrored in numerous field studies of clinical physicians (Lusted, 1977; DeSmet *et al.*, 1979; Christensen-Szalanski & Bushyhead, 1981). In contrast, studies of weather forecasters (Winkler & Murphy, 1968; Murphy & Winkler, 1971, 1974, 1977) and "experts" who set odds for horse racing (Dowie, 1976) show the experts to have generally excellent calibration. Lichtenstein *et al.* (1982) explain the difference between the results for physicians and weather forecasters by three factors. First, weather forecasters are used to making probabilistic forecasts. Second, forecasting the weather is reasonably repetitive, while diagnosing the ailments of patients is highly non-repetitive. Third, the outcome feedback for weather forecasters is reasonably prompt and well understood. In contrast, physicians often will not know for quite a long time after, if their diagnoses have been correct. This could be due to patients taking their time to report back, going elsewhere or the link between ailment and symptom not being well-defined. These results suggest that feedback and task structure are important determinants of an expert's calibration capabilities.

To summarise, on the various factors which might be expected to condition the relationship between confidence and accuracy, it is only for the "difficulty effect" where the bulk of the research findings tend to point in an unambiguous direction, namely, the more difficult the task the more likely is a subject to be over-confident. Despite the ambiguity of the conclusions for the other factors described above, it seems clear that they must be borne in mind when examining issues which deal with the relationship between confidence and accuracy.

In addition to the research findings just described, there is some existing empirical work in forecasting which suggests that a confidence based consensus measure might be more highly correlated with accuracy than the traditional consensus measure. Dalkey (1969) used the confidence of participants in their responses to reach more accurate estimates of quantitative phenomena in Delphi exercises. Dalkey found that by selecting for inclusion in the feedback only those responses considered "highly confident" by their proponents, a slightly superior result was achieved. Dalkey *et al.* (1970) found that in situations in which the "highly confident" group was reasonably large, a definitely superior result could be achieved.

Of course, in reality the relationship between accuracy and consensus (however defined) is complex because it will depend upon the level and range of accuracy achieved by experts, the degree of agreement across the decisions and the correctness of decision-makers' confidences. Thus, how a confidence based consensus measure compares to the conventional consensus measure is an empirical issue.

A new measure of consensus

The experiment used to compare a confidence based consensus measure to the conventional consensus measure is concerned with failure prediction. Because of the dichotomous nature of the predictions made (fail, not fail) Ashton's absolute measure of consensus, defined as the number of cases a pair of subjects agree in their predictions, is used to represent the conventional consensus measure. In order to facilitate comparison with this conventional consensus measure, the form of the new pairwise consensus measure to be used here is defined as follows:

(i) if a pair of experts, j and $j+1$, agree on prediction i, then the new pairwise consensus measure for the pair $(j, j+1)$ for prediction i will equal their average level of confidence, i.e.

$$(\text{confidence}_{j,i} + \text{confidence}_{j+1,i})/2;$$

(ii) if a pair of experts, j and $j+1$, disagree on prediction i, then the new pairwise consensus measure for the pair $(j,j+1)$ for prediction i will equal -1 times their average level of confidence, i.e.

$$-1^{*}(\text{confidence}_{j,i} + \text{confidence}_{j+1,i})/2.$$

This form of consensus produces a pairwise measure for each prediction which can range from -100% (when both subjects are 100% confident in their decisions but disagree) to $+100\%$ (when both subjects are 100% confident and agree).

It might be argued that this new measure of consensus has an inherent ambiguity, in that it gives the same score of zero to pairs of subjects which:

(i) agree on prediction but have zero confidence; or which

(ii) disagree on prediction and have zero confidence.

The authors would argue however, that there is no ambiguity here because (i) and (ii) are the same judgement. This is because an expert would only give a zero confidence weighting to a decision which s/he believed to be random. There are, of course, no substantial differences between agreement and disagreement over random decisions.

SUBJECTS AND THE STUDY TASK

The experiment used in this paper was also designed to tackle other issues, namely, whether small company account users were able to use post-1981 Companies Act "abridged accounts" successfully and whether they were able to utilise additional, bank reference, information when making predictions. Thus, since the article dealing with these issues has already been published (Keasey & Watson, 1986), the details of the experiment will only be briefly described here.

The experiment was concerned with failure/non-failure predictions and was similar to the one conducted by Kida (1980), the results of which were used by Ashton in her study of the relationship between accuracy and consensus. In terms of conducting the experiment, arrangements were made with the organisers of the 1985 U.K. Trade Credit Conference for the authors to give a paper on risk assessment methods based on an evaluation of the delegates' success at handling a number of "risk assessment" case studies. Details of proposed conference sessions were included in a circular sent to all members of the Institute of Credit Management during the spring of 1985 and delegates were asked to nominate the sessions they would attend at the forthcoming conference. Eight-two delegates elected to attend the authors' session and were duly sent, during the early summer of 1985, a package of risk assessment case studies and asked to reply as soon as possible. By early September, 36 complete and usable replies and a further 22 incomplete/apologies/excuses had been received. Early and late respondents were analysed for statistically significant differences in responses. None were found and the full set of 36 subjects was used for the analysis described in the following sections.

Each trade credit specialist was provided with three year financial profiles for 20 (10 failed and 10 non-failed) real but disguised small (£10,000–£100,000 total assets), single plant, independently owned mechanical engineering companies operating in the north-east of England. The participants were informed that the companies had the above the characteristics and were told that companies of this type had a 50% chance of failing, where failure was defined as ceased trading within three years of the last set of accounts. The 50% probability of failure over a three-year period is close to the actual base rate for this type of company (see Storey *et al.*, 1987).

Each financial profile was based upon the post-1981 Companies Act "abridged accounts" of the companies. These abridged financial statements, which are the only accounting information increasingly large numbers of small companies actually lodge at Companies House, consist only of a balance sheet and accompanying notes. From these statements five financial ratios, for each of the three years prior to failure, were chosen to reflect, as far as was possible given the limitations of these financial statements, the major categories of financial information and common usage/reference in the accounting and finance literatures.

For each of the 20 companies, the subjects were asked to estimate on the basis of the information provided whether the company would fail within three years of the last set of accounts. Subjects were also asked to indicate their degree of confidence for each prediction on a 100 point scale. To check whether subjects had used the confidence scale in a consistent manner, at the end of the exercise they were asked to state how confident they were, on a 100 point scale, in their overall performance. At this point the subjects were also asked to state how long the task had taken them, their age, their years of experience and how they felt about the information provided given the nature of the task.

To determine whether the accuracy, consensus and confidence of the trade credit specialists' decisions were affected by the cue set provided, half of the initial sample of 82 subjects were supplied with additional information in the form of bank references detailing the bank's assessment of the financial condition of the company. Bank references were chosen as the additional cue after discussions with the Institute of Credit Management. These discussions indicated that the assessment provided by a company's bankers were one of the most important cues used by trade credit specialists. Therefore, the addition of bank references seemed the obvious way of increasing the non-redundant, experimental cue set.

Of the 36 subjects who completed the study, 18 had been supplied profiles without the bank reference information. There were no significant differences (see Keasey & Watson, 1986) between the bank reference group and the no-bank reference group in terms of age, experience, time taken to complete the task and in their comments concerning how useful they found the information provided.

Thus, the major differences of the present experiment as compared to the one used by Ashton are, in terms of the size of companies (small as compared to large), the profession (trade credit specialists rather than auditors) and nationality (being U.K. rather than U.S. based) of the participants, the incorporation of subjects' confidence in their predictions to develop the new form of consensus and the effect upon accuracy, consensus and confidence of increasing the cue set provided via the control group method.

ANALYSIS AND RESULTS

Due to the dichotomous nature of the predictions made by the subjects in this study, only Ashton's absolute measures of consensus and accuracy and the new consensus measure are used in the analysis. Both pairwise and individual measures of accuracy and consensus will be used in the following analysis to allow comparison across the arguments of interest. For as argued by Ashton (1985)

> "the traditional measures of consensus and accuracy are not comparable. Consensus traditionally is expressed as agreement between pairs of subjects while accuracy typically is measured for an individual subject. Consequently, to facilitate comparisons between these two dimensions of judgement performance, new measures of accuracy and consensus were developed" (p. 174).

Pairwise absolute accuracy is defined as the mean accuracy score over the 20 cases for each pair of subjects. Individual absolute accuracy is the number of cases (out of 20) that a subject made the correct failure/non-failure prediction. The pairwise measure for conventional consensus is the number of cases each pair of individuals agree in their predictions. The pairwise measure for the new confidence based consensus is as defined in the previous section. Individual consensus, for both measures and for both subgroups of participants (bank and no-bank), is the mean consensus over the 17 pairs in which a participant is included.

Summary statistics and t*-tests*

Table 1 presents the summary statistics for the three measures of interest for both types of analysis. The summary statistics for the no-bank reference group and the bank reference group indicate that the provision of the additional cue of a bank reference does not significantly affect accuracy but significantly reduces both measures of consensus. In terms of explaining this result, the existing evidence from psychology is not particularly helpful. If the additional cue is redundant or irrelevant, consensus decreases (Einhorn, 1971; Well, 1971; Nystedt, 1974). However, if the cue is relevant and uncorrelated with prior information, consensus can either decrease (Nystedt & Magnusson, 1972; Nystedt, 1974) or increase (Huff & Friedman, 1967). As it is difficult to say whether bank reference information was perceived as being irrelevant or not by the subjects, the only conclusion to be drawn is that the additional bank reference cue decreases consensus.

Accuracy for the no-bank reference group also shows that the subjects are less accurate than those analysed by Ashton — the group mean for correct predictions being 66% as compared to the 84% mean achieved by the auditors in Ashton's study. To test for the "environmental predictability" of the financial information provided, a stepwise linear discriminant model succeeded in correctly classifying 75% (70% and 80% of failed and non-failed companies, respectively) of the 20 cases. Whilst this level of predictive accuracy is low compared to the results obtained in the majority of large company studies, the model's accuracy is greater than could be expected from random choice. The major reason for the low level of predictive accuracy being the inherent variability of small firm performance (Storey *et al.*, 1987) and the well recognised unreliability of their financial reporting (ASC, 1979, 1980, 1983).

TABLE 1. Summary statistics and *t*-tests

	No-bank group		Bank group		
	Mean	S.D.	Mean	S.D.	*t*-value
Pairwise analysis *N* = 153 (in each group)					
Accuracy	13.33	1.20	13.17	1.28	1.19
Conventional consensus	15.02	2.67	13.04	3.04	6.02*
New consensus	30.99	15.51	19.06	18.24	6.17*
Individual analysis *N* = 18 (in each group)					
Accuracy	13.33	1.78	13.17	1.92	0.27
Conventional consensus	15.02	1.46	13.04	1.60	3.88*
New consensus	30.99	9.93	19.06	10.09	3.58*

*Significant at 0.01 level.

Correlation analysis

Table 2 shows the correlation coefficients between the various arguments for each subgroup. For the no-bank reference group, the Pearson (Spearman) correlation between accuracy and the conventional consensus measure, is positive and significant at a 1% level for the pairwise analysis and at a 5% (10%) level for the individual analysis. Thus, the positive correlations reported by Ashton, whilst also evident in this

TABLE 2. Correlation analysis (Pearson/Spearman in parentheses)

	No-bank group				Bank group			
Pairwise analysis $N = 153$ (in each group)								
Accuracy	0.47‡	(0.35)‡			0.58‡	(0.56)‡		
Conventional consensus	0.96‡	(0.93)‡	0.38‡	(0.24)‡	0.96‡	(0.94)‡	0.52‡	(0.50)‡
	New consensus		Accuracy		New consensus		Accuracy	
Individual analysis $N = 18$ (in each group)								
Accuracy	0.52‡	(0.41)†			0.73‡	(0.51)‡		
Conventional consensus	0.96‡	(0.84)‡	0.43†	(0.33)•	0.96‡	(0.93)‡	0.70‡	(0.38)†
	New consensus		Accuracy		New consensus		Accuracy	
(One-tailed) significance tests for differences in correlation coefficients for the two consensus measures and accuracy								
Pairwise	$Z = 0.96$	$P = 0.16$			$Z = 0.78$	$P = 0.22$		
Individual	$Z = 0.30$	$P = 0.38$			$Z = 0.19$	$P = 0.42$		

• ≤ 10%
† ≤ 5%
‡ ≤ 1%

study, are somewhat reduced. Also, the positive relationship between accuracy and consensus increases when the cue set is increased. Ashton's results, therefore, appear to be corroborated by the results presented here.

Turning to the central hypothesis of this paper, the results of Table 2, indicate that the new measure of consensus is, for both groups and for both types of analysis, more highly correlated with accuracy than the conventional consensus measure. However, as indicated by the "Fisher's transformation of *r*" *z*-statistics shown in the table (see Glass & Stanley, 1970), the differences in correlation coefficients are not statistically significant at conventional levels of confidence.

Thus, for the group of experts and task considered here, the new measure of consensus does not appear to provide a markedly better alternative than the conventional consensus measure. This does not mean, however, that the new consensus measure does not provide information which might be considered complementary to that provided by the conventional consensus measure. To test whether the new consensus measure provided incremental information as compared to conventional consensus, a two stage regression approach was adopted. The first stage derived uncorrelated residuals from a regression of the new consensus measure on the conventional consensus measure. The second stage regressed the accuracy scores on the uncorrelated residuals from the first stage. This procedure provides semi-partial correlations between the new consensus measure and accuracy, which measure the unique contribution of the new consensus measure after the contribution of the conventional consensus measure has been "partialled out". The semi-partial correlations for the no-bank ($sr = 0.37$) and bank ($sr = 0.29$) pairwise analysis groups are significant at 1% levels of confidence. For the individual analysis, however, only the semi-partial correlation of 0.39 for the no-bank group was significant at 5% levels of confidence (sr for bank group = 0.22).

The above analysis indicates that the new measure of consensus provides statistically significant additional information in respect of explaining subjects' accuracy. This suggests, for the present experiment at least, that the two measures ought not to be considered as either competitors or substitutes. Rather, they should be viewed as complements since they jointly succeed in explaining a statistically significantly greater proportion of the variance in subjects' accuracy than the conventional consensus measure alone.

CONCLUDING REMARKS

In this paper the role assigned to consensus within experimental studies of decision-making has been briefly considered and a new measure of consensus which incorporates experts' confidence in their decisions has been proposed. The experimental evidence presented in this paper shows that, whilst the new confidence based consensus measure is not significantly more highly correlated with accuracy than the traditional consensus measure, it is complementary to the conventional consensus measure. The single set of results presented in this paper should, of course, not be seen on their own as offering conclusive evidence for the adoption/non-adoption of the new measure. Instead they should be seen as an initial set of results from a single set of experts.

The preceding analysis also develops the understanding of the relationship between accuracy and consensus in the following ways. First, in a more "difficult" decision setting than those considered by Ashton, the relationship between accuracy and consensus (however defined) is reassuringly significantly positive, if to a somewhat lesser extent than reported by Ashton. Second, although the actual level of agreement between individuals and pairs of individuals significantly falls when the available cue set is increased, the positive correlation between accuracy and consensus appears to be strengthened. More work dealing with this relationship and the redundancy/non-redundancy of cue changes is clearly needed. Finally, the findings suggest that much more work is needed to understand how the new measure of consensus might perform across a range of decision situations which control for the difficulty of the task and expertise of the subjects.

BIBLIOGRAPHY

Adams, P. A. & Adams, J. K., Training in Confidence Judgments, *American Journal of Psychology* (1958) pp. 747–751.

Adams, J. K. & Adams, P. A., Realism of Confidence Judgements, *Psychological Review* (1961) pp. 33–45.

Ashton, A. H., Does Consensus Imply Accuracy in Accounting Studies of Decision-making, *The Accounting Review* (April 1985) pp. 173–185.

Auditing Practices Committee, *Small Companies: The Need for Audit* (ICAEW, 1979).

Auditing Practices Committee, *Analysis of Public Comment on Small Companies: The Need for Audit* (ICAEW, 1980).

Auditing Practices Committee, *Auditing Standards Guidelines* (ICAEW, 1983).

Bar-Hillel, M., The Base Rate Fallacy in Probability Judgements, *Acta Psychologica* (1980) pp. 221–233.

Clarke, F. R., Confidence Rates, Second-choice Responses, and Confusion Matrices in Intelligibility Tests, *Journal of the Acoustical Society of America* (1960) pp. 35–46.

Choo, G. T. G., Training and Generalisation in Assessing Probabilities for Discrete Events, technical report 76-5, Brunel Institute of Organizational and Social Studies, Uxbridge, England (1976).

Christensen-Szalanski, J. J. J. & Busheyhead, J. B., Physicians Use of Probabilistic Information in a Real Clinical Setting, *Journal of Experimental Psychology: Human Perception and Performance* (1981) pp. 928–935.

Dalkey, N. C., The Delphi Method: An Experimental Study of Group Opinion, working paper 5888-PR, Rand Corp. (1969).

Dalkey, N. C., Brown, B. & Cochran, S., The Delphi Method IV: Effect of Percentile Feedback and Feed-in of Relevant Facts, working paper 6118-PR, Rand Corp. (1970).

DeSmet, A. A., Fryback, D. G. & Thornbury, J. R., A Second Look at the Utility of Radiographic Skill Examination of Trauma, *American Journal of Radiology* (1979) pp. 95–99.

Dowie, J., On the Efficiency and Equity of Betting Markets, *Economica* (1976) pp. 139–150.

Einhorn, H. J., Use of Non-linear, Non-compensatory Models as a Function of Task and Amount of Information, *Organizational Behavior and Human Performance* (1971) pp. 1–27.

Einhorn, H. J., Learning from Experience and Suboptimal Rules in Decision-making, in Kahneman, D., Slovic,

P. and Tversky, A. (eds) *Judgement under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982).

Glass, G. & Stanley, J. C., *Statistical Methods in Education and Psychology* (Prentice-Hall, 1970).

Hazard, T. H. & Peterson, C. R., Odds Versus Probabilities for Categorical Events, technical report 73-2, Decisions and Designs, Inc., McLean, VA (1973).

Huff, F. W. & Friedman, H., A Study of the Effects of Agreement Among Clinicians of Redundant and New Information, Confidence and Time Available for Assessments, *Journal of General Psychology* (1967).

Keasey, K. & Watson, R., The Prediction of Small Company Failure: Some Behavioural Evidence for the U.K., *Accounting and Business Research* (Winter 1986) pp. 49–57.

Kida, T., An Investigation into Auditors' Continuity and Related Qualification Judgements, *Journal of Accounting Research* (Autumn 1980) pp. 506–523.

Koriat, A., Lichtenstein, S. & Fischhoff, B., Reasons for Confidence, *Journal of Experimental Psychology: Human Learning and Memory* (1980) pp. 107–118.

Libby, R. & Lewis, B. L., Human Information Processing Research in Accounting: The State of the Art in 1982, *Accounting, Organizations and Society* (1982) pp. 231–285.

Lichtenstein, S. & Fischhoff, B., Do Those Who Know More Also Know More About How Much They Know? The Calibration of Probability Judgements, *Organizational Behaviour and Human Performance* (1977) pp. 159–183.

Lichtenstein, S. & Fischhoff, B., How Well do Probability Experts Assess Probability?, decision research report 80-5, Decision Research, Eugene OR (1980a).

Lichtenstein, S. & Fischhoff, B., Training for Calibration, *Organizational Behaviour and Human Performance* (1980b) pp. 149–171.

Lichtenstein, S., Fischhoff, B. & Phillips, L. D., Calibration of Probabilities: The State of the Art to 1980, in Kahneman, D., Slovic, P. and Tversky, A. (eds) *Judgement Under Uncertainty: Heuristics and Biases* (Cambridge University Press, 1982).

Lusted, L. B., A Study of the Efficacy of Diagnostic Radiologic Procedures: Final Report on Diagnostic Efficiency, Efficacy Study Committee of the American College of Radiology, Chicago (1977).

Nickerson, R. S. & McGoldrick, C. C., Confidence Ratings and the Level of Performance on a Judgemental Task, *Perceptual Motor Skills* (1965) pp. 311–316.

Murphy, A. H. & Winkler, R. L., Forecasters and Probability Forecasts: Some Current Problems, *Bulletin of the American Meteorological Society* (1971) pp. 239–247.

Murphy, A. H. & Winkler, R. L., Subjective Probability Forecasting Experiments in Meteorology: Some Preliminary Results, *Bulletin of the American Meteorological Society* (1974) pp. 1206–1216.

Murphy, A. H. & Winkler, R. L., Can Weather Forecasters Formulate Reliable Probability Forecasts of Precipitation and Temperature?, *National Weather Digest* (1977) pp. 2–9.

Nystedt, L., Consensus Among Judges as a Function of Amount of Information, *Education and Psychological Measurement* (1974).

Nystedt, L. & Magnusson, D., Predictive Efficiency as a Function of Amount of Information, *Multivariate Behavioural Research* (1972).

Oskamp, S., Over Confidence in Case Study Judgements, *Journal of Consulting Psychology* (1965) pp. 261–265.

Pitz, G. F., Subjective Probability Distributions for Imperfectly Known Quantities, in Gregg, L. W. (ed.), *Knowledge and Cognition* (New York: Wiley, 1974).

Sieber, J. E., Effects of Decision Importance on Ability to Generate Warranted Subjective Uncertainty, *Journal of Personality and Social Psychology* (1974) pp. 688–694.

Storey, D. J., Keasey, K., Watson, R. & Wynarczyk, P., *The Performance of Small Firms* (Croom-Helm, 1987).

Tversky, A. & Kahneman, D., Judgement under Uncertainty: Heuristics and Biases, *Science* (September 1974) pp. 1124–1131.

Weil, A. D., The Influence of Irrelevant Information on Speeded Classification Tasks, *Perceptions and Psychophysics* (1971).

Winkler, R. L. & Murphy, A. H., "Good" Probability Assessors, *Journal of Applied Meteorology* (1968) pp. 751–758.

Accounting, Organizations and Society, Vol. 14, No. 4, pp. 347–357, 1989.
Printed in Great Britain

0361–3682/89 $3.00+.00
Pergamon Press plc

THE IMPACT OF PARTICIPATION IN BUDGETING AND JOB DIFFICULTY ON MANAGERIAL PERFORMANCE AND WORK MOTIVATION: A RESEARCH NOTE*

LOKMAN MIA
University of Tasmania

Abstract

This study examined the interactive effect of participation and job difficulty on managerial performance and work motivation in a budgetary context. Using middle-level managers from six companies working in various functional areas as subjects, this study found that managerial performance was high where perceived participation was commensurate with perceived level of job difficulty. On the other hand, performance was low where participation was not commensurate with the perceived level of job difficulty. No significant interactions were obtained with respect to work motivation.

The argument that participation in budgeting may be positively related to managerial performance and motivation is based on the rationale that when a goal or a standard set by a participative process is accepted, employees: (a) internalize the established goal or standard; and (b) feel a personal responsibility to achieve it (Milani, 1972). Consequently, employees develop greater motivation and hence attain better performance. Researchers also argue that participation may increase employees' personal feelings of "ownership" of the goal or standard and the level of their trust in the intentions of management (Vroom, 1964; Lowin, 1968; Patchen, 1970; Lawler & Hackman, 1979; Raval, 1982). Others have argued that participation is a mechanism for information exchange by which individual employees obtain more information about their job (Hopwood, 1976). Such additional information facilitates a clearer understanding of their job, thereby improving motivation and performance (Mitchell, 1973; Schuler & Kim, 1976; Schuler, 1980).

Nevertheless, there is only equivocal evidence supporting the posited relationships between participation and performance and motivation. While a number of studies report a significant positive relationship between participation and performance in a budgetary context (Argyris, 1952; Becker & Green, 1962; Brownell, 1979; Kenis, 1979), other studies did not observe such a relationship (French *et al.*, 1960; Fleishman, 1965; Milani, 1975; Ivancevich, 1976; Steers, 1976). Moreover, some studies have reported a negative relationship between the two (Bryan & Locke, 1967; Stedry, 1960). Similarly, equivocal results have been obtained with respect to the posited participation–motivation relationship in a budgetary context. While Vroom (1964) and Hofstede (1967), for example, reported a significant positive relationship between participation in budgeting and motivation, Brownell & McInnes (1986) did not observe such a relationship (see also Staw, 1977; Ivancevich, 1979).

One possible reason for the equivocal results

---

*I am indebted to Peter Brownell, Macquarie University, Ken Ferris, Southern Methodist University and Ron Weber and Peter Ziegler, Queensland University for their comments on an earlier version of the paper. Comments by participants in seminars on the paper at University of Otago, University of New South Wales and at Canterbury University are appreciated. I am also indebted to the anonymous reviewers for their useful comments and suggestions on earlier drafts of the paper.

reported in previous studies is that the relationships may be moderated by contingency (situational) variables (see Brownell, 1981). Hopwood (1976), for example, suggests that the moderating role of contingency variables, such as job characteristics, personality characteristics and social conditions, be considered while studying participation in budgeting. A review of the literature reveals that a number of studies have investigated the relationship between budgetary participation and employee behaviour while considering the contingent role of personality variables. Brownell (1981), for example, reported that locus of control (a personality variable) moderated the relationship between job satisfaction, performance and budgetary participation. Previously, Vroom (1959) and Hofstede (1967) reported that personality variables such as authoritarianism and need for independence moderated the relationship between motivation, performance and participation in decision-making.

No studies, however, have addressed the possible moderating role of a job characteristic such as job difficulty on the relationship between budgetary participation and employee behaviour, particularly at the individual level (i.e. using individual employees as the unit of analysis). This study, conducted in a field setting, investigates such a moderating role by examining how job difficulty interacts with participation to influence managerial performance and motivation.

The remainder of this section discusses the moderating effects of job difficulty and develops the hypotheses to be tested. Subsequent sections focus on the method of the study, the results and a discussion of the results, limitations and conclusions.

JOB DIFFICULTY AS A MODERATING VARIABLE

Two important characteristics of a job or task (henceforth, job) are its difficulty and variability (Van de Ven & Delbecq, 1974). A job may be difficult because of its complexity, heterogeneity, unpredictability, lack of a prescribed performance procedure or because of its changing operational technology (Perrow, 1967). A job may also be difficult because of its variability. Indeed, Van de Ven & Delbecq (1974) argue that when a job becomes less difficult as a result of learning, it is likely to become less variable because a body of knowledge may have been developed specifying a prescribed way to perform it. Routineness is another important characteristic of a job. A routine job may be programmed by the prescription of rules and procedures which facilitate performance (Aitken, 1960; Fullan, 1970) and, as a consequence, the job becomes less difficult. Thus, difficulty is an important job characteristic and it may be a surrogate for other important job characteristics such as variability and routineness.

A potential moderating effect of job difficulty on the participation–performance and participation–motivation relationships appears to be supported by the extant literature. In a high job difficulty situation, more information is required at the point of task execution (Fry & Slocum, 1984). Managers operating in such a situation may not have all the necessary information to perform the job; they may need to obtain and process additional information to understand the job clearly, thereby improving their chances of successful performance and obtaining desired rewards.

Participation as a mechanism for information exchange may enable managers to obtain and process additional information about their job (Hopwood, 1976; Galbraith, 1977). Such additional information may facilitate understanding of their job requirements, thereby increasing their confidence (expectancy) that efforts made to perform the job may be successful. Under expectancy theory of motivation (see Vroom, 1959; House, 1971; Lawler & Suttle, 1977; Ferris, 1977, 1978), an increase in the expectancy that effort leads to successful performance (E->P) may increase managers' motivation to perform their job. Thus, it can be argued that participation in budgeting may improve managers' motivation and performance particularly in a high job difficulty situation. Hopwood (1976, pp. 83–85), for example, argues that in the case of a high job difficulty situation, participation in budgeting may offer greater benefits than a non-participative style.

In a low job difficulty situation, on the other hand, it is possible to develop and apply predetermined rules, policies and standards for the performance of a job (Tushman & Nadler, 1978). Thus, individual employees operating under these conditions may be able to perform the job only by knowing the prescribed rules, policies and standards; they may not need to obtain and process additional information for this purpose and, therefore, have less need for participation. If work is routine (less difficult), most decisions involving the task are likely to be routine hence, managers will be more motivated and perform better if their time is not wasted by participation in decisions with obvious solutions (see Strauss, 1982).

Based on the discussions above, it can be argued that the relationships between participation and performance and between participation and motivation are contingent upon an appropriate fit between participation and job difficulty. A high(low) level of participation under a high(low) job difficulty situation (the appropriate fit) is likely to have a more positive (strong) relationship with managerial performance and motivation. The converse is expected when the appropriate fit does not occur (see Fig. 1). The hypotheses stating this proposition are as follows:

> H_1. The higher the perceived job difficulty (X_2), the more positive is the relationship between participation in budgeting (X_1) and managerial performance (Y). In mathematical terms, dY/dX_1 is greater when X_2 increases.
>
> H_2. The higher the perceived job difficulty (X_2), the more positive is the relationship between participation in budgeting (X_1) and motivation (M). In mathematical terms, dM/dX_1 is greater when X_2 increases.

The hypotheses are summarised in Fig. 1.

METHODS

Research site and sample

Data for the study were collected using a survey questionnaire from middle-level managers working in six companies operating in New Zealand. The number of people employed by each of these companies was between 100 and 1500. Of these six companies, one manufactures food items, one makes furniture and wood products, one assembles automobiles, two provide retail services and one provides banking services.

Discussions in preliminary meetings with the managements of these companies indicated that, for several years, these companies have been: (a) using comprehensive budgeting systems; and (b) interested in encouraging a participative style of budgeting in all departments/sections. Moreover, it was also evident from the discussions that: (i) employees needed time to understand and follow various policies of the company; and (ii) an employee holding at least a middle-level management position had adequate opportunities to participate in setting and implementing his or her (henceforth, his)

		Job Difficulty	
		High	Low
Participation	High	More Positive Relationship	Less Positive Relationship
	Low	Less Positive Relationship	More Positive Relationship

Fig. 1. The model of the study.

departmental/sectional budgets and goals. The two criteria for selecting participants were:

(1) the individual had been working for the company for at least six months prior to the study; and

(2) the individual was holding a management position and was responsible for preparing and/or implementing the budget or target for his department/section.

The managements of the six companies provided 93 participants from among the managers who were available at the time of the study and who met the two selection criteria.[1]

Each participant was sent a questionnaire together with a letter explaining the objective of the study and assuring anonymity of his responses. Each participant was provided with a self-addressed prepaid envelope to return the questionnaire directly to the researcher.

Of the 93 managers selected for the study, 76 fully completed and returned the questionnaire yielding a response rate of 81.7%. These managers were from production (20%), sales (16%), marketing (6%), accounting (18%), import and purchasing (13%), financing (10%), personnel (10%) and MIS (7%). On average, the respondents had been employed for five and a half years by their respective companies and supervised, on average, 13 subordinates.

Measurement of variables

The questionnaire obtained measures for three variables: participation in budgeting, job difficulty and work motivation.

Participation in budgeting. Participation refers to the amount of influence and involvement that an individual employee perceives he or she has on a jointly-set budget (Milani, 1975; Brownell, 1979). Milani's participation continuum scale was used to assess the employee's perceived amount of participation in budgeting. Since Milani's measure has been widely used, the results of this study can be integrated into the existing body of literature (Brownell, 1982). Examples of items included in Milani's scale are: "Which category below best describes your activity when the budget is being set? I am involved in setting:" and "How much influence do you feel you have on the final budget?". The managers rated their level of perceived participation in budgeting under each of the 6 items on a 7-point Likert-type scale anchored at both ends. A manager's overall score for participation was the sum of scores for the 6 items. A check of internal reliability yielded a Cronbach (1951) alpha coefficient of 0.91 for this scale. Moreover, a factor analysis of the scores yielded one eigenvalue greater than unity, hence, one factor was extracted suggesting that the scale was unidimensional. The factor accounted for 69.9% of the variance. Table 1 presents the factor loadings, eigenvalue and the percentage of explained variance.

Job difficulty. This variable refers to the degree of complexity of the search process in performing a job, the amount of thinking time required to solve job-related problems and the body of knowledge that provides guidelines for performing the job (Perrow, 1967). Job difficulty was measured using a 7-item Likert-type scale developed by Van de Ven & Delbecq (1974). Examples of items included in this measure are: "To what extent is there an under-

---

[1] The company managements agreed to select the participants for the study using the two selection criteria only. This was achieved by making the managements fully aware of the rationale for the selection criteria and the importance of using those criteria in selecting the participants. It is believed that participants having different personality characteristics had a fair chance to be included in the sample. Thus, the research design is expected to be *randomized for personality characteristics*. One may also argue that since the participants had been working for their company for a fairly long time (five and a half years on the average), they might have become used to the particular style of management in the company. Consequently, the possible influence of their individual personality characteristics on the effectiveness of their participation in budgeting might have become neutral. Since data for this study were collected from companies operating in New Zealand only, the results may be specific to New Zealand culture.

TABLE 1. Factor analysis of participation and job difficulty responses

Participation		Job difficulty	
Items	Loadings *factor 1*	Items	Loadings *factor 1*
1	0.86	1	0.59
2	0.73	2	0.55
3	0.71	3	0.61
4	0.90	4	0.63
5	0.89	5	0.72
6	0.94	6	0.65
		7	0.70
Eigenvalue	4.19	Eigenvalue	2.85
Variance explained	70%	Variance explained	41%

standable sequence of steps that can be followed in doing your work?" and "On average, how much actual ***thinking time*** do you usually spend trying to solve such specific difficult problems?". Managers rated their perceived level of job difficulty on a 7-point scale anchored at both ends. A manager's overall score for job difficulty was the sum of his or her scores for the 7 items. A check of internal reliability yielded a Cronbach (1951) alpha coefficient of 0.75 for this scale. Moreover, a factor analysis of the scores yielded one eigenvalue greater than unity, hence, one factor was extracted suggesting that the measure was unidimensional. The factor accounted for 41% of the variance. Table 1 presents the factor loadings, the eigenvalue, and the percentage of explained variance.

Work motivation. An employee's work motivation refers to a joint function of his or her expectancy that effort (E) towards a job performance (P) will lead to the successful performance of the job (i.e. $E->P$), his or her expectancy that successful performance of the job will lead to attainment of some outcomes (i.e. $P->O$) and his or her valence (V) or desirability of those outcomes (Ferris, 1977). The construct was measured using the scale developed by Lawler & Suttle (1973). Part one of the questionnaire consisted of 20 expectancy items: 2 items measured ($E->P$) expectancies and 18 items measured ($P->O$) expectancies. For ($E->P$) expectancies, effort was represented by the expression "working hard", for ($P->O$) expectancies, performance was represented by expressions such as "high productivity" and outcome by expressions such as "promotion". Each of the expectancy items called for a response on a 7-point Likert-type scale ranging from never (1) to always (7). Part two of the questionnaire consisted of 20 items measuring the valence (desirability) of the 20 task outcomes. Respondents rated their valence on a 9-point Likert-type scale, ranging from extremely undesirable (1) to extremely desirable (9). The 20 outcomes included both intrinsic and extrinsic rewards. The overall score for a manager's motivation was determined by multiplicatively combining his or her scores under the expectancy and valence items.[2] The Lawler & Suttle's instrument has been shown to be reliable, and it has been widely used (Ferris, 1977). An internal reliability check yielded Cronbach (1951) alpha coefficients of 0.71 for the measure of ($E->P$) expectancy, 0.82 for ($P->O$) expectancy and 0.72 for valence (V); these results compare favourably with previous results[3] (Ferris, 1977).

---

[2] Work motivation: $M=(E->P)\Sigma[(P->O)(V)]$ where ($E->p$) = effort leads to performance expentancy, ($P->O$) = performance leads to outcome expectancy and (v) = valence or desirability of outcome or reward. This study used a between-subjects analysis of motivation scores. An alternative to this method would be a within-subjects analysis (Johnson, 1979). However, following Ferris (1977, 1979) and Jiambalvo (1979), a between-subjects analysis method was used here for two reasons. First, this method is comparatively easy to use and understand. Jiambalvo (1979, p. 448), for example, argues that "Research using a within-individuals analysis becomes cumbersome very quickly". Second, empirical evidence shows that a within-subjects analysis is not superior to a between-subjects analysis. Examples of studies supporting this argument are Parker & Dyer (1976), Kopelman (1979) and Murray & Frazier (1986) (see also Ferris, 1979). It is expected that individual employees have a desire to obtain both intrinsic and extrinsic rewards (intrinsic and extrinsic valences) for, say, good job performance. Consequently, treating the two valences separately in constructing motivation scores would result in an incomplete explanation for an individual's motivation. Moreover, the relationship, if any, between employees' participation and their extrinsic valences is not theoretically clear (see Ronen & Livingstone, 1975; Brownell & McInnes, 1986).

Managerial performance. Performance in this context refers to the degree of successful role achievement attained by an employee (Ferris, 1977). Although previous studies of budget participation have used self-ratings (Brownell, 1979, 1982; Merchant, 1981), the current study used superiors' ratings of performance for two reasons. First, self-ratings are likely to have a higher leniency error (higher mean values) and a lower variability error (restricted range) in the score (see Prien & Liske, 1962; Thornton, 1968). More specifically, the likelihood that a manager will rate his or her own unsatisfactory (below average) performance as unsatisfactory is at least low if not nil. Second, the superiors' ratings are likely to provide a wider (more global) evaluation of managerial performance (Heneman, 1974).

The list of participants provided by management of the six companies included the name and address of each participating manager's superior and the period for which the participant had been working under his or her current superior. At least a period of three months was considered necessary for a superior to form an accurate opinion about his or her subordinate's performance. Five of the 76 managers (who fully completed and returned the questionnaire) did not meet the above criterion. The superior of each of the remaining 71 managers was requested to accurately evaluate the overall performance of the manager reporting directly to him or her using a 9-point Likert-type scale ranging from (1) well below average to (9) well above average performance.

Each of the respondent superiors was assured of anonymity and informed that the information collected would be used only for the purpose of the study.[4] Superiors' ratings for 62 managers were obtained, hence the final sample size was 62 for the analysis of the performance data and 76 for the motivation data.

RESULTS

Table 2 presents descriptive statistics and Table 3 the correlation matrix for the variables utilized in this study. As there is no significant multicollinearity amongst the variables as shown in Table 3 (see Nie *et al.*, 1975, p. 340), regression analysis was used to test hypotheses H_1 and H_2 (Blalock, 1972; Schoonhoven, 1981). Moreover, the multiplier interaction model presented in equation (1) is efficient in testing the

TABLE 2. Descriptive statistics for the variables

Variables	Means	S.D.	Actual range	Theoretical range
Performance	5.15	2.82	2–9	1–9
Motivation	39.06	16.90	0–90	na
Participation	29.12	8.71	9–40	6–42
Job difficulty	21.26	4.97	10–33	7–35

TABLE 3. Correlation matrix for the variables

		Y	M	X_1	X_2	X_1X_2
Y	Performance	1.00				
M	Motivation	0.10	1.00			
X_1	Participation	0.18	0.16	1.00		
X_2	Job difficulty	−0.21*	0.01	0.02	1.00	
X_1X_2	Interaction	0.31†	0.03	−0.08	−0.23*	1.00

* $p<0.05$.
† $p<0.01$.

---

[3] Since the measure of motivation involves a multiplicative model (see note 2), a factor analysis of the responses would not be useful.

[4] The superiors were also requested to evaluate performance of a participating manager only if they (superiors) could do it without any personal prejudice. Superiors of six participating managers refused to evaluate their (manager's) performance on personal grounds.

hypothesized interaction effect (Khandwalla, 1974; Allison, 1977; Brownell, 1981, 1982; Schoonhoven, 1981; Govindarajan, 1986; Chenhall, 1986).

Hypothesis H_1 was tested using the regression model:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_1X_2 + e... \quad (1)$$

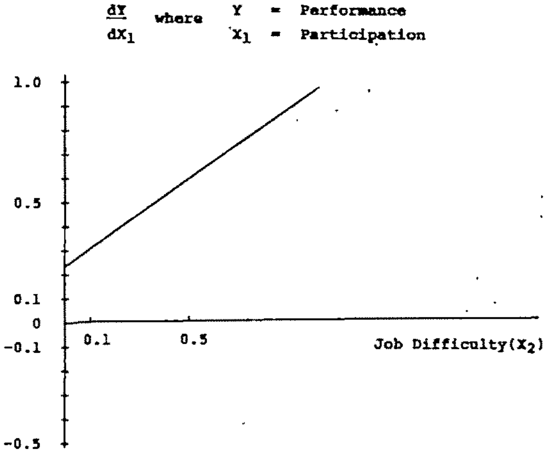
where Y is managerial performance, X_1, participation in budgeting, X_2, job difficulty and X_1X_2 is interaction.

To permit an acceptance of hypothesis H_1, the coefficient b_3 in equation (1) is required to be: (a) significant, indicating a significant interaction between participation and job difficulty affecting performance; and (b) positive to support the direction of the hypothesis.

The results presented in Table 4 reveal that the coefficient b_3 was positive and significant ($p < 0.04$). The prediction model in equation (1) explained 13.83% of the variance ($R^2 = 13.83\%$ (adjusted $R^2 = 9.4\%$), $F = 3.10, p < 0.03$) in performance. Therefore, the results indicate that budgetary participation and job difficulty interacted at a significant level to positively impact managerial performance. Thus, H_1 was accepted.

Figure 2 presents a graphical representation of the results to enable a clearer understanding of the significant interaction between budgetary participation and job difficulty in relation to managerial performance. Drawing the figure involved taking the partial derivative of equation (1) over X_1 and then plotting the joint effect of the main and interaction terms (see Southwood, 1978; Schoonhoven, 1981; Govindarajan, 1986). Figure 2 indicates that the higher the job difficulty, the greater the positive impact of participation on performance (see Schoonhoven, 1981, p. 364).

Hypothesis H_2 was tested using a model identical to equation (1) by substituting the motiva-



where $\frac{dY}{dX_1} = b_1 + b_3X_2$

from $Y = a + b_1X_1 + b_2X_2 + b_3X_1X_2 + e$

Fig. 2. Significant interaction between participation and job difficulty affecting performance.

TABLE 4. Test of hypothesis H_1.
The interaction between participation and job difficulty affecting performance

Variables		Coefficient	T^* value	Level of significance (p)
X_1	Participation	0.24	1.22	ns
X_2	Job difficulty	−0.39	1.16	ns
X_1X_2	Interaction	0.69	2.05	<0.04
	Constant	6.41	39.27	<0.001

$R^{2\dagger} = 13.83\%, F(3,58) = 3.10, p<0.03.$
$n = 62.$

*One-tailed test.
†Adjusted $R^2 = 9.4\%$.

tion (M) data for performance (P) data. The results were not significant and, therefore, hypothesis H_2 was rejected.[5] An implication of the results is discussed later in the paper.

DISCUSSION, LIMITATIONS AND CONCLUSIONS

As expected, the proposed interaction between participation and job difficulty affecting performance was significant. Performance was found to be high when the amount of participation was proportionate to the level of job difficulty. In contrast, performance was found to be low when the amount of participation was disproportionate to the level of job difficulty. Thus, a congruence (good fit) between the level of participation and job difficulty appeared to be a pre-condition for effective participation (in terms of performance).

These results support the argument that employees may find participation useful when they perceive their job difficulty to be high and when their perceived level of job difficulty is low, they may find participation relatively ineffective. This study, therefore, offers one explanation for the equivocal results of previous studies which reported a negative or no relationship between participation and performance (Blumenfield & Leidy, 1969; Milani, 1975; Bryan & Locke, 1976; Kenis, 1979; Steers, 1979). The level of participation in budgeting by respondents in those studies might have been disproportionate to the level of perceived job difficulty. Similarly, this study offers one explanation for the results of previous studies which reported a positive relationship between participation and performance (Schuler & Kim, 1979; Schuler, 1980; Merchant, 1981; Brownell, 1982); the level of participation in budgeting by respondents in those studies might have been proportionate to the level of job difficulty they perceived.

The results for work motivation are disappointing but not surprising. Research results for motivation have been equivocal. While some studies (e.g. Patchen, 1970; Siegal and Ruh, 1973; Collins, 1978; Kenis, 1979) have reported a significant positive relationship between participation and motivation, other studies (e.g. Staw, 1977; Ivancevich, 1979; Brownell, 1983) observe no such relationship. More recently, Brownell & McInnes (1986) could not obtain a significant positive relationship between participation and motivation. One possible explanation for such equivocal results is the apparent lack of a consistent approach to measuring motivation in previous studies investigating the participation–motivation relationship (see Brownell & McInnes, 1986). The approach to defining and measuring motivation in the current study is consistent with that taken by Brownell & McInnes (1986). Both these studies used an expectancy model for defining, and the scale developed by Lawler & Suttle (1973) for measuring motivation. The insignificant results for motivation obtained in the current study are therefore consistent with those reported by Brownell & McInnes (1986). Furthermore, based on the results of this study, it can be argued that perceived job difficulty does not act as a contingent variable to moderate the participation–motivation relationship. Thus, the results of the study attempt to improve our understanding of the lack of a relationship between participation and motivation.

With respect to the limitations of the present study, several points are worth mentioning. First, the managements of the six companies selected the participants for the study and, hence, the sample was not strictly random. Consequently, some caution is necessary before generalizing the results. Second, the regression model of performance (equation (1)) explained a relatively low level of variance ($R^2 = 13.8\%$) indicating that variables other than budgetary participation and job difficulty are at work affecting managerial performance. Future research needs to pursue this issue further.

---

[5] Hypothesis H_2 was tested using an alternative formulation of motivation, i.e. $M=\Sigma[(E\text{->}O)(V)]$ as suggested by Ferris (1977, p. 606). The results were not significant and hence, rejection of hypothesis H_2 was further confirmed.

Notwithstanding such limitations, the results of this study might help to reconcile the results reported by previous studies of participation in budgeting. The study therefore makes some contribution to improving our understanding of the effect of budgetary context. For practising managers, the results suggest that participation in budgeting as a tool for improving managerial performance should be used selectively. In cases where perceived job difficulty is high, encouraging managers to participate in budgeting is likely to improve their performance. On the other hand, in cases where perceived job difficulty is low, encouraging managers to participate in budgeting is unlikely to improve their performance. In such cases, prescribed rules, programmes, policies and standards may be more effective in the pursuit of improved performance.

BIBLIOGRAPHY

Allison, P. D., Testing for Interaction in Multiple Regression, *American Journal of Sociology* (1977) pp. 144–153.

Aitken, H. G. J., *Taylorism at Watertown Arsenal* (Cambridge, MA: Harvard University, 1960).

Argyis, C., *The Impact of Budgeting on People* (New York: Controllership Foundation, Inc., 1952).

Becker, S. & Green, D., Budgeting and Employee Behavior, *Journal of Business* (1962) pp. 392–402.

Blalock, H. M. Jr, *Social Statistics*, 2nd edn (New York: McGraw-Hill, 1972).

Blumenfield, W. S. & Leidy, T. R., Effectiveness of Goal Setting as a Management Device: Research Note, *Psychological Reports* (1969) p. 752.

Brownell, P., Participation in Budgeting, Locus of Control and Organizational Effectiveness, unpublished PhD. dissertation, (University of California, Berkeley, 1979).

Brownell, P., Participation in Budgeting, Locus of Control and Organizational Effectiveness, *The Accounting Review* (1981) pp. 844–860.

Brownell, P., The Role of Accounting Data in Performance Evaluation, Budgetary Participation and Organizational Effectiveness, *Journal of Accounting Research* (1982) pp. 12–27.

Brownell, P., The Motivational Impact of Management-By-Exception in a Budgetary Context, *The Journal of Accounting Research* (1983) pp. 456–472.

Brownell, P. & McInnes, J. M., Budgetary Participation, Motivation and Managerial Performance, *The Accounting Review* (1986) pp. 587–600.

Bryan, J. F. & Locke, E. A., Goal Setting as a Means of Increasing Motivation, *Journal of Applied Psychology* (1967) pp. 274–277.

Collins, F., The Interaction of Budget Characteristics and Personality Variables with Budgetary Attitudes, *The Accounting Review* (1978) pp. 324–335.

Chenhall, R. H., Authoritarianism and Participative Budgeting: A Dyadic Analysis, *The Accounting Review* (1986) pp. 263–272.

Cronbach, L. T., Coefficient Alpha and the Internal Structure of Tests, *Psychometrica* (1951) pp. 297–334.

Ferris, K. R., A Test of the Expectancy Theory of Motivation in an Accounting Environment, *The Accounting Review* (1977) pp. 605–615.

Ferris, K. R., A Test of the Expectancy Theory of Motivation in an Accounting Environment: A Response, *The Accounting Review* (1979) pp. 412–413.

Fleishman, E. A., Attitude Versus Skill Factors in Work Group Productivity, *Personnel Psychology* (1965) pp. 253–266.

French, J. R. P., Israel, J. & Aas, D., An Experiment on Participation in a Norwegian Factory, *Human Relations* (1960) pp. 3–20.

Fry, L. W. & Slocum, J. W. Jr, Technology, Structure and Workgroup Effectiveness: A Test of a Contingency Model, *Academy of Management Journal* (1984) pp. 221–246.

Fullan, M., Industrial Technology and Worker Integration in the Organization, *American Sociological Review* (1970) pp. 1028–1039.

Galbraith, J. R., *Organization Design* (Reading, MA: Addison-Wesley, 1977).

Govindarajan, V., Impact of Participation in the Budgetary Process on Managerial Attitude and Performance: Universalistic and Contingency Perspectives, *Decision Sciences* (1986) pp. 496–516.

Heneman, H. G., Comparisons of Self- and Superior Ratings of Managerial Performance, *Journal of Applied Psychology* (1974) pp. 638–642.

Hofstede, G. H., *The Game of Budget Control* (The Netherlands: Koniklijke Van Gorcum and Company, N.V. Assen, 1967).

Hopwood, A., *Accounting and Human Behaviour* (Englewood Cliffs, NJ: Prentice-Hall, 1976).

House, R. J., A Path-Goal Theory of Leader Effectiveness, *Administrative Science Quarterly* (1971) pp. 321–338.

Ivancevich, J. M., Effects of Goal Setting on Performance and Job Satisfaction, *Journal of Applied Psychology* (1976) pp. 605–612.

Ivancevich, J. M., An Analysis of Participation in Decision Making Among Project Engineers, *Academy of Management Journal* (1979) pp. 253–269.

Jiambalvo, J., Performance Evaluation and Directed Job Effort: Model Development and Analysis in a CPA Firm Setting, *Journal of Accounting Research* (1979) pp. 436–455.

Johnson, W. B., A Test of the Expectancy Theory of Motivation in an Accounting Environment: A Commentary, *The Accounting Review* (1979) pp. 409–411.

Khandwalla, P. M., Mass Output Orientation of Operations Technology and Organizational Structure, *Administrative Science Quarterly* (1974) pp. 74–79.

Kenis, I., Effects of Budgetary Goal Characteristics on Managerial Attitudes and Performance, *The Accounting Review* (1979) pp. 707–721.

Kopelman, R. E., Across-individual, Within-individual and Return on Effort Version of Expectancy Theory, *Decision Sciences* (1979) pp. 651–652.

Lawler, E. E. & Hackman, J. R., Impact of Employee Participation in the Development of Pay Incentive Plans: A Field Experiment, *Journal of Applied Psychology* (1979) pp. 467–471.

Lawler, E. E. & Suttle, L. R., Expectancy Theory and Job Behaviour, *Organizational Behaviour and Human Performance* (1973) pp. 482–583.

Lowin, A., Participative Decision-making: A Model, Literature Critique, and Prescriptions for Research, *Organizational Behaviour and Human Performance* (1968) pp. 68–106.

Merchant, K. A., The Deisgn of the Corporate Budgeting System: Influences on Managerial Behavior and Performance, *The Accounting Review* (1981) pp. 813–829.

Milani, K. R., Budgeting in an Industrial Setting: A Field Study of the Relationship Between Involvement in Budget-setting and Job Performance and Attitudes, unpublished Ph.D. Thesis, The University of Iowa (1972).

Milani, K. R., The Relationship of Participation in Budget-setting to Industrial Supervisor Performance and Attitudes: A Field Study, *Accounting Review* (1975) pp. 274–284.

Mitchell, T. R., Motivation and Participation: An Integration, *Academy of Management Journal* (1973) pp. 274–284.

Murray, D. & Frazier, K. B., A Within-subjects Test of Expectancy Theory in a Public Accounting Environment, *Journal of Accounting Research* (1986) pp. 400–404.

Parker, D. F. & Dyer, L., Expectancy Theory as a Within-person Behavioral Choice Model: An Emprical Test of Some Conceptual and Methodological Refinements, *Organizational Behaviour and Human Performance* (1976) pp. 97–117.

Patchen, M., *Participation, Achievement, and Involvement on the Job* (Englewood Cliffs, NJ: Prentice-Hall, 1970).

Perrow, C., *Organizational Analysis: A Sociological Perspective* (Belmont, CA: Wadsworth, 1967).

Prien, E. P. & Likse, R. E., Assessment of Higher-level Personnel: III. A Comparative Analysis of Supervisor Ratings and Incumbent Self-ratings of Job Performance, *Personnel Psychology* (1962) pp. 187–194.

Raval, V., Differences in Meaning and Relationship Between Levels of Participation in Budget Setting, *American Institute for Decision Sciences, 14th Annual Meeting* (1982) pp. 37–38.

Ronen, J. & Livingstone, L., An Expectancy Theory Approach to the Motivational Impact of Budgeting, *The Accounting Review* (1975) pp. 671–685.

Schuler, R. S., A Role Expectancy Perception Model of Participation in Decision Making, *Academy of Management Journal* (1980) pp. 311–340.

Schuler, R. S. & Kim, J. S., Interactive Effect of Participation in Decision Making, the Goal Setting Process and Feedback on Employee Satisfaction and Performance, *Academy of Management Proceedings* (1976) pp. 114–117.

Schoonhoven, C. B., Problems with Contingency Theory: Testing Assumptions Hidden Within the Language of Contingency Theory, *Administrative Science Quarterly* (1981) pp. 349–377.

Siegel, A. L. & Ruh, R. A., Job Involvement, Participation in Decision Making, Personal Background, and Job Behavior, *Organizational Behaviour and Human Performance* (1973) pp. 218–327.

Southwood, K. E., Substantive Theory and Statistical Interaction: Five Models, *American Journal of Sociology* (1978) pp. 1154–1203.

Staw, B. M., Motivation in Organization: Towards Synthesis and Redirection in Staw, B. M. and Salancik, G. R. (eds) *New Directions in Organizational Behavior* (Chicago: St. Clair Press, 1977).

Stedry, A. C., *Budget Control and Cost Behavior*, (Englewood Cliffs, NJ: Prentice-Hall, 1960).

Steers, R. M., Factors Affecting Job Attitudes in a Goal Setting Environment, *Academy of Management Journal* (1976) pp. 6–16.

Strauss, G., Workers Participation in Management: An International Perspective in Staw, B. and Cummings, L. (eds) *Research in Organization Behavior* pp. 173–266 (Greenwich, CT: JAI Press, 1982).

Thornton, G. C., The Relationship Between Supervisory and Self-appraisals of Executive Performance, *Personnel Psychology* (1968) pp. 441–456.

Tushman, M. L. & Nadler, D. A., Information Processing as an Integrating Concept in Organization Design, *Academy of Management Review* (1978) pp. 613–624.

Van de Ven, A. H. & Delbecq, A. L., A Task Contingent Model of Work Unit Structure, *Administrative Science Quarterly* (1974) pp. 183–197.

Vroom, V. H., Some Personality Determinants of the Effects of Participation (Stanford University, 1959).

Vroom, V. H., *Work and Motivation* (New York: John Wiley, 1964).

Accounting, Organizations and Society, Vol. 14, No. 4, pp. 359–378, 1989.
Printed in Great Britain

0361–3682/89 $3.00+.00
Pergamon Press plc

ASSESSING INHERENT RISK DURING AUDIT PLANNING: THE DEVELOPMENT OF A KNOWLEDGE BASED MODEL*

JAMES M. PETERS
College of Business Administration, University of Oregon

BARRY L. LEWIS
College of Business and Administration, University of Colorado, Boulder

and

VASANT DHAR
Department of Computer Applications, New York University

Abstract

Within the academic and professional auditing communities there has been growing concern with accurately assessing the various risks associated with the performance of an audit. One approach to developing sophisticated risk assessment models is to study how experienced expert auditors use industry and firm specific factors in making audit judgments. This study presents a model of inherent risk assessment based on literature reviews and a field study that involved structured and unstructured interviews and observations of experts in audit planning meetings. Analysis of the data gathered led to the specification of a conceptual model of inherent risk assessment which has been implemented as a computer program (a computational model). Auditors were asked to assess the behavior and performance of the computational model as a first step in evaluating the expert model.

Over the last five years there has been a growing interest in using computer based cognitive models to study the audit process. The state of the art of these research efforts has been reviewed by Messier & Hansen (1986). This paper describes progress to date on a research project to develop conceptual and computational (computer program) models of inherent risk assessment in audit planning. The project is based on an analysis of the various phases involved in a complete audit process as opposed to isolated audit tasks. This global perspective makes it more likely that specific models developed as a part of this research will integrate readily into a comprehensive model of the entire audit process.

Operationalizing the conceptual model of inherent risk in terms of a computational model has several advantages for studying and modeling audit decision-making. First, the computer program is a precise and testable expression of a human information processing model. Another advantage is that computer programs also allow for comprehensive testing of a model under different conditions of the task environment in that they can be run with data sets that vary in controlled ways. In addition, the program can be

---

*Funding for this project was provided by the Peat Marwick Foundation through their Research Opportunities in Auditing Program (Project Number 84-118) and by a Faculty Research Grant of the Graduate School of Business of the University of Pittsburgh. We are especially grateful for the many hours contributed by the professional accountants who formed the basis for our conceptual model of risk analysis. We have benefited from the useful comments of Terry O'Keefe and two anonymous referees.

built to provide traces of its own behavior. In this way, patterns of behavior can be studied and compared to human behavior under controlled conditions. A pragmatic benefit of developing a computer model is the possibility of using such a model to support human judgment. If the program incorporates the expertise of experienced auditors and has access to an accumulating store of client data, it can be used to make audit judgments for validating or augmenting judgments of auditors on specific cases.

This research project has resulted in the development of conceptual and computational models of inherent risk assessment in auditing. Inherent risk assessment was identified as an appropriate starting point in the overall audit process after a series of loosely structured interviews with auditors and a review of professional and academic literature. Analysis of the data gathered in these initial efforts led us to conclude that audit planning constitutes an *ex-ante* rehearsal of the entire audit process. Within planning, inherent risk assessment was identified as the central activity involving judgment that forms the basis for audit planning.

The remainder of this paper is divided into five sections. First, we discuss the concept of inherent risk assessment in auditing by reviewing the extant literature. In the third section we discuss the observation techniques we have used in attempting to understand the audit process and, more specifically, inherent risk evaluation. This is followed by a description of the conceptual model of inherent risk assessment. The fifth section presents the computational model that embodies the basic features of the conceptual model. The sixth section discusses auditors' reactions to the initial model. We conclude with a summary of the current state of our project and future research that we intend to conduct as part of this project.

INHERENT RISK ASSESSMENT

An important first step to developing a model of inherent risk assessment is to determine the requirements it places on the auditor. We began by reviewing the extant literature to determine what these task requirements were. The literature can be classified into two main groups: prescriptive (i.e. how auditors should perform inherent risk assessment) and descriptive (i.e. how auditors actually perform inherent risk assessment). The prescriptive literature includes pronouncements of the AICPA and audit firm manuals. The descriptive literature includes several academic studies.

Prescriptive literature

Inherent risk assessment is part of a larger problem of audit risk assessment. Both these concepts have been discussed widely in the professional and academic literature. Cushing & Loebbecke (1983) provide a detailed discussion of the historical development of the risk model, a review of the literature dealing with the appropriateness of the model and a discussion of the problems that auditors might encounter in trying to implement it. The majority of this literature has focused on the AICPA's audit risk model as presented in SAS 47:

$$AR = IR * CR * DR,$$

where AR is audit risk, IR is inherent risk, CR is control risk and DR is detection risk.

Inherent risk is the susceptibility of an account balance or class of account balances to error that could be material assuming that there are no related internal accounting controls. Control risk is the risk that this error in the account balance will not be caught by the client's internal control system. Detection risk is the risk that any error not detected by the control system will not be detected by the audit procedures, thereby affecting financial statements. The AICPA emphasizes that the above model of audit risk is a purely conceptual one, and says little about whether or how its various components can be measured.

According to the AICPA, the risk assessment process should occur during audit planning. Auditors should determine an acceptable level for audit risk, assess the levels of inherent risk and control risk and then determine the level of detection risk based on these assessments. Auditors should assess the inherent riskiness of specific accounts by reviewing a variety of factors that are specific to the client, the client's industry or the economy in general and by determining the impact of these factors on individual accounts. The purpose of this assessment is to help determine the nature, timing and extent of tests of the client's internal control systems and transactions (AICPA, 1985).

The audit approaches of several Big Eight CPA firms call for the assessment of inherent risk at two levels during the audit planning process: the engagement level and the account level (Arthur Andersen and Co., 1983; Elliott, 1983; Grobstein & Craig, 1984; Peat Marwick International, 1985). At the engagement level, inherent risk assessment is used to consider whether the client is too risky to retain and to provide some perspective on acceptable levels of individual account risk. At the account level the main concern is assessing the degree of risk associated with a given account balance in the general ledger. The primary emphasis of these approaches is on assessing risk at the individual account level.

The professional literature also provides lists of factors that should be considered in assessing inherent risk. Peat Marwick International's (1985) audit manual, for example, presents a representative list:

(1) monetary amount associated with the account;

(2) susceptibility of asset to theft;

(3) complexity required to determine amounts to be entered in the account;

(4) degree of management judgment involved in valuing the account;

(5) degree to which external events affect values in the account;

(6) past history of error;

(7) degree to which client's financial condition motivates management to misstate the amount in the account; and

(8) experience of the personnel performing accounting functions involving the account.

Descriptive approaches

The descriptive approaches to inherent risk assessment in the academic literature have attempted to examine the cognitive ability of auditors to implement the risk model and to determined what factors auditors believe to be important in the assessment of inherent risk.

Cognitive abilities. Boritz *et al* (1986) note that the risk evaluation task as characterized by the risk model can be represented as a problem of decision-making under risk. Their benchmark for cognitive ability in such a task is the subjective expected utility (SEU) model. Based on an extensive discussion of the audit judgment literature, they conclude that there is substantial evidence that auditors do not conform to the SEU model when performing these judgments.

The result of two other studies provide conflicting opinions of the auditor's ability to make the kinds of judgments required by the audit risk model. Jiambalvo & Waller (1984) concluded that "... auditor's intuitive combination of the risk componets did not correspond closely with the kind of combination dictated by the audit risk model of SAS 39". However, Libby *et al.* (1985) concluded that "[i]t is refreshing that auditor subjects' performance so closely matched predictions from the audit risk model." A major limitation of audit judgment research is that the tasks that auditors are asked to perform often bear little relationship to the tasks that they perform in actual practice. Since judgments may be sensitive to even small changes in the task environment (Einhorn & Hogarth, 1981), the use of unrealistic judgment tasks could explain many of the divergent findings in the literature.

Factors that affect inherent risk. Boritz *et al.* (1986) gave their subjects four case studies for which ten audit planning judgments had been made and asked the subjects to review and evaluate these judgments, to provide a revised set of judgments based on their evaluation and to rate the difficulty of each judgment. One of the ten judgments was an assessment of the inherent riskiness of the client. They found that inherent risk assessments were sensitive to liquidity problems, inventory levels, strength of internal controls, history of errors and fraud and whether the audit is connected with an initial public offering. They concluded that inherent risk judgments have a broader interpretation than that implied by the definitions provided in professional standards. They indicated that one reason for this broader interpretation by their subjects might be that the subjects were asked to assess the inherent riskiness of the audit whereas such assessments are usually based on individual accounts.

Gibbins & Wolf (1982) used an extensive questionnaire to elicit from auditors sets of items that were the best predictors of whether significant audit problems would be encountered for a hypothetical audit. The auditors were also asked to list other items that were highly related to the predictive items. The auditors were given a set of items from which to choose but were also allowed to add items of their own. The Gibbins & Wolf study dealt with five stages of the audit process of which planning was one. The responses that relate to the planning stage are most relevant to this research. The items that were identified most often can be classified as either client factors or audit firm factors. Client factors include the quality of the accounting system and internal controls, the competence of top management, the accounting staff and the internal audit staff, plans for major financing activities, financial condition and profitability and the client's direct interest in the conduct of the audit. Audit firm considerations include availability and continuity of staff resources, fee considerations, audit approach and the competence of the audit manager and field staff.

Although the Gibbins & Wolf task was not to specifically assess inherent risk but to identify factors that could contribute to potential problems in the audit, the two task are similar enough so that their results can be used to provide some insight into the inherent risk assessment process.

Summary of risk model implications

The conclusions most relevant to this research that can be drawn from the audit risk literature involve the identification of factors that affect inherent risk and the professional guidance given to auditors for assessing inherent risk. These conclusions can be summarized as follows.

(1) A review of the factors that affect inherent risk as presented in audit manuals and as determined in the descriptive studies discussed earlier indicates that the factors considered by auditors are broad and include not only client factors but also audit firm factors. Some of the factors considered are easily quantifiable (e.g. client's liquidity position) while others are not (e.g. quality of client management or client wishes with respect to the conduct of the audit).

(2) A review of the factor lists that were developed in the descriptive studies indicates that inherent risk is not independent of control risk in that auditors consider some general assessment of control risk when assessing inherent risk.

(3) The results of descriptive studies also imply auditors use some method for selecting, evaluating and combining the diverse sets of factors mentioned above in coming up with judgments on inherent risk. What is not clear, however, is the relationship among these factors and how these relationships are used intuitively by experienced auditors in assessing inherent risk.

RESEARCH METHOD

The long range purpose of this research effort is to develop a theory of how auditors assess inherent risk in a normal audit environment, to test the theory by developing a computational model that reflects the theory and to compare the computational model's behavior to that of experienced auditors.

Research projects that use cognitive computer models to test theories of human judgment normally proceed in several distinct phases. First, the judgment task and its environment are investigated by reviewing relevant literature and by interviewing experts in the judgment task. Second, more sharply focused data are gathered to refine the researcher's understanding of the task and its environment and to develop an initial theory of how the task is performed. The methods used in this second phase are varied but usually include structured interviews and/or concurrent verbal protocols based on case data.

In the third phase, a theory is derived from the data and a computational model built that reflects the theory. Developing a computational model of expert behavior requires a detailed understanding of the *knowledge components* involved in problem solving, and a mapping of these components into a representation scheme that can be used to generate expert behavior. In effect, the validation of the computer model is achieved by subjecting it to stimuli that a human expert might be expected to deal with and observing whether the *process* and the *outcome* of executing the model match those of the expert. The real bottleneck in developing the model is typically the lack of adequate and correct problem solving data from the expert.

The fourth phase attempts to overcome this bottleneck. Experts are asked to interact with the computational model and to evaluate its performance. Their evaluation of both process and outcome provides additional problem solving data which can be used to revise the computational model (and often the underlying theory). This is an iterative process and can be quite extensive and span several years. Once the computational model's performance has reached a satisfactory level, it it tested in more controlled ways. This final phase usually involves having experts provide blind evaluations of the model's behavior and/or comparing the model's behavior on a given case to that of experts evaluating the same case.

The research efforts reported in this paper have followed this general model through the beginning of the fourth phase. We have developed the initial theory (conceptual model) as well as an initial computational model and have received preliminary evaluations from experienced auditors. Our initial data were obtained through archival sources and unstructured interviews with several auditors. The main limitation of interviews, of course, is that they can produce "summary behavior" which may not be a correct manifestation of the problem solving process which one might observe with real cases (Dhar & Pople, 1987). Nevertheless, the interviews were useful in several respects. Specifically, they helped reveal the following.

(1) A delineation of various phases involved in the overall audit process. These phases consist of planning, tests of the control system, substantive testing, evaluation of evidence and formulation of an opinion.

(2) Aspects of the audit process that auditors find particularly difficult (such as identifying specific characteristics of general ledger accounts that might increase the risk associated with an account).

(3) Areas of the audit in which expert judgment is required.

(4) The level of detail of the data involved in the various phases of the audit process.

The level of understanding of the audit problem gained through the interviews set the stage for subsequent, more targeted data gathering. Specifically, the first part of the data gathering process made it clear that it would be best too focus first on the audit *planning* phase since this phase results in a context within which subsequent judgments about a client are made. In particular, the concept of inherent risk evaluation was identified as being central to determining how to plan an audit in the first place. For this reason, subsequent data gathering focused on inherent risk evaluation as performed in actual audit situations.

Several methods were employed in the second stage of inquiry. These approaches were all designed to enable us to identify the factors that auditors felt were important in assessing inherent risk and to identify the process used to analyze inherent risk. First, we asked auditors to review the documentation of the planning activity for audits on which they had recently worked and to attempt to recall the data and reasoning they used to arrive at their conclusions concerning inherent risk. Second, we arranged a role playing exercise in which one audit manager pretended he was turning over one of his clients to another audit manager as part of a normal rotation of staff. Finally, we observed unobtrusively two actual audit planning meetings. The observations were followed immediately by a brief question and answer session with each participant. Except for one of the planning meetings, all of the data collection sessions were tape recorded and transcribed for later analysis.[1] Field notes were taken during the last planning session and a complete analysis of the session was produced immediately after the meeting and question and answer sessions were over.

We elected not to fashion realistic cases to present to the expert for solution, a method often used by cognitive psychologists and AI researchers for inducing and observing real expert problem solving behavior because of the lack of a well established information processing model of inherent risk assessment. Without the clear guidance of such a model, we were concerned that any cases we developed might not accurately capture the task environment and therefore distort our theory development.

The analysis of the data also proceeded in several phases. The first phase consisted of scanning the transcripts and field notes looking for any statement that would identify a factor that had influenced an inherent risk assessment. That is, statements of the general form "I altered the scope on that account because..." or "I felt that that account would be more risky because...". Analysis of the resulting list of factors produced a simple classification scheme that grouped the factors into three main classes: general firm factors that affected the overall riskiness of the firm but were not directly associated with a given account, account factors that could directly affect a given account and audit firm factors. Table 1 contains a representative sample of the factors identified and the classification scheme developed.

The list of factors and the classification scheme provided evidence concerning the knowledge components involved in inherent risk assessment and the level of fineness or specificity with which auditors dealt with these components. The second phase of data analysis consisted of reanalyzing the transcripts and field notes to determine why the factors in the list were important and how auditors made linkages between these factors and levels of inherent risk. The specific procedure involved locating places in the text where inherent risk factors were mentioned and reviewing the surrounding text for any explicit comments about why the factor was important and how it affected risk. Such explicit comments were rare. Consequently, the bulk of our analysis in this phase involved generating hypotheses about the relationship of factors to risk based on our general business knowledge and confirming these hypotheses with the auditors.

For example, one subject stated that he had been a little more skeptical about management's judgments, particularly related to valuation accounts, when he discovered a potential management buy-out plan. He went on to state that the audit scope for the valuation accounts (e.g. allowance for doubtful accounts) had been increased in this case. We inferred from his comments that:

(1) the auditor felt management motives were important in assessing risk because of the importance of management judgment in determining and therefore potentially manipulating, certain

TABLE 1. Classification of audit risk factors

	External to firm	Within the firm
General firm factors	Nature of the industry	Level of management integrity
Direct account factors	Loss of major customer	High turnover of key personnel
Audit firm factors	Amount of turnover in audit team	Not applicable

---

[1] After reading the transcript from the first planning session, the participating firm decided the information was too sensitive to record. They did consent to our taking notes during the second session. We felt that the importance of having access to real data in real situations justified the extra effort and potential data interpretation problems that are associated with field notes.

account balances; and

(2) the auditor was using knowledge about the effects changes in valuation accounts would have on reported earnings, the relationship between reported earnings and stock prices and the effect stock prices would have on the potential cost of a management buy-out as a basis for developing his concern for the increased risk association with valuation accounts.

Comments of this type were important in helping us develop the conceptual model and the assertions supporting it which are presented in the next section. In addition, an analysis of the inferred relationships led us to include a general model of fundamental economic relationships that exist within any firm as part of computational model which is described later.

CONCEPTUAL MODEL

This section presents a description of our conceptual model of inherent risk evaluation. An overview of the conceptual model is provided, a list of specific assertions implied by the model are discussed and examples of evidence that support those assertions are presented. We relate these assertions in terms of a process model that is designed to reflect the actual process of inherent risk assessment. The declarative knowledge used in the process model is described in the next section where we discuss the computational model.

Since much of our analysis described in the previous section was subjective, we present raw transcript excerpts in this section to allow the reader to draw their own conclusions concerning the evidence supporting the assertions. Figure 1 presents a graphic representation of the process model and is intended to provide a global view of the inherent risk evaluation process.

General description of the process model

The process model reflects what we term a *differential approach* towards assessing the inherent risk of material error in a given general ledger account balance. More specifically, our model asserts that the auditor begins the inherent risk evaluation process by generating expectations for accounts balances. The auditor identifies changes that have occurred in the firm or its environment and determines how those changes should interact with historic trends to produce an expected balance in the account. In order to do this, we posit that the auditor uses an understanding of the relationships between firm/environmental factors and general ledger accounts. We term these relationships as the "internal model of the firm" (described in the next section). By using this model of the firm, the auditor determines how the changes should affect the balances in a given account. For accounts that contain actual balances that are outside a reasonable range of the expected balance, the auditor reviews factors that might create or affect management incentives to misstate the account balances (e.g. the existence of a compensation plan keyed to reported earnings). At the same time, the auditor considers factors that might impact on the likelihood that management could or would misstate that particular account balance (e.g. the degree of judgment allowed in the determination of account balances). The auditor would also consider such unintentional factors as the complexity of the transactions or accounting for a particular account since such factors might also be responsible for deviations from expected balances. Based on this analysis, the auditor decides if additional evidence will be needed to determine whether the difference between the expected balance and the actual balance was due to an error in the expectation generating process, a legitimate response by management to a change in the environment, an unintentional error or a questionable response by management.

Empirical support for the model

The inherent risk evaluation process described above is expectation driven, where expectations about general ledger account balances are based on knowledge of historical trends and specific changes in the audited firm's environment that affect the audited firm. Further, the output of this evaluation process is

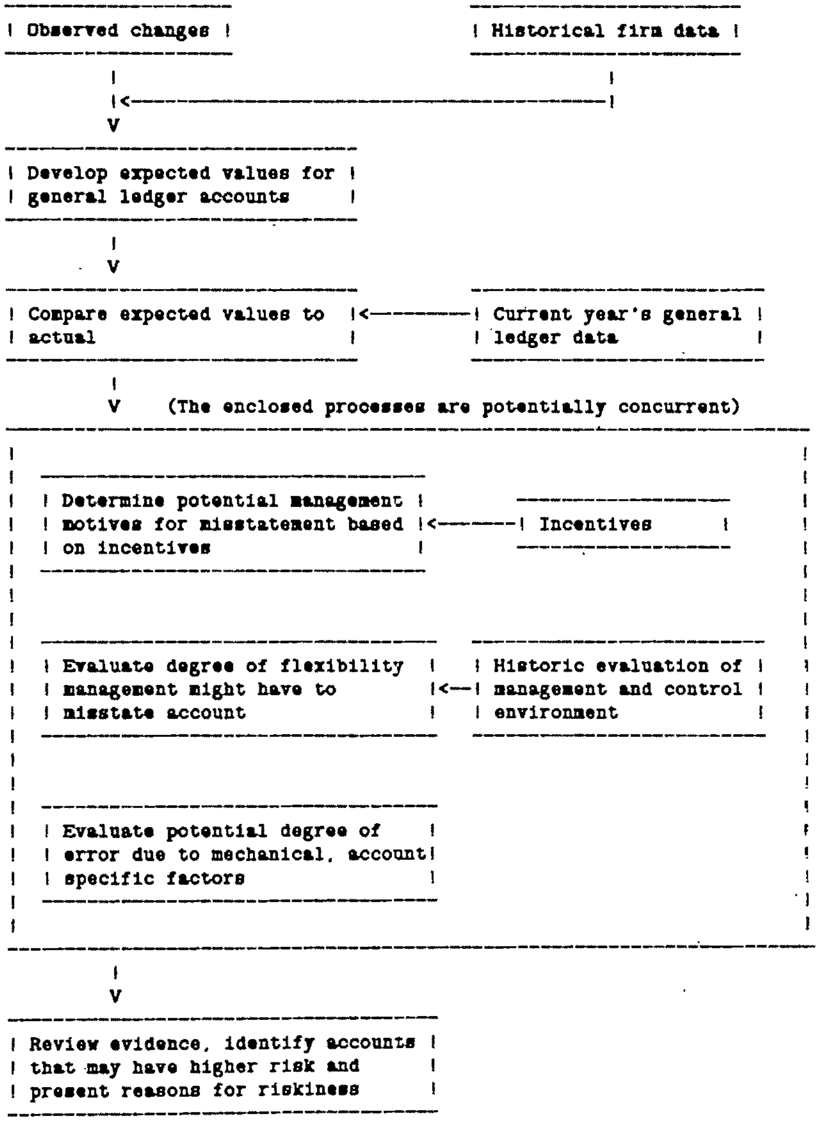

Observed changes
Historical firm data
Develop expected values for general ledger accounts
Compare expected values to actual
Current year's general ledger data
(The enclosed processes are potentially concurrent)
Determine potential management motives for misstatement based on incentives
Incentives
Evaluate degree of flexibility management might have to misstate account
Historic evaluation of management and control environment
Evaluate potential degree of error due to mechanical, account specific factors
Review evidence, identify accounts that may have higher risk and present reasons for riskiness

Fig. 1. Conceptual model of inherent risk evaluation.

an analysis of factors that might explain differences between expected and actual account balances. These factors include management's incentives and abilities to affect the account balance as well as the complexity associated with calculating the account's balance.

In the remainder of this section, we provide examples of transcript evidence that support the assertions of our conceptual model. Although we offer only a few representative samples, we should note that all of our transcripts contain statements that support all of the assertions in the model. The fact that the transcripts were taken from data collection session that involved different firms, different auditors at different levels of responsibility and in different task settings provides strong support for the model. Appendix A contains a short version of each assertion and a sampling of taped quotations supporting that assertion. The following discussion presents a general summary of the evidence for each assertion. In the section describing the computational model we discuss how specifically these assertions are used.

Assertion 1: inherent risk assessments are generated on an account by account basis. There was strong evidence across all transcripts that auditors ultimately boil risk assessment down to the account level in order to be more precise in adjusting their audit effort. Additional support for this assertion was found in the audit procedures of the two audit firms participating in our study. One of the CPA firms used an optional form that records auditors' assessments of inherent risk on an account by account basis. The other firm was beginning to use a form that tied account level risk assessment to the use of specific audit procedures which address those risks.

Assertion 2: auditors generate expectations concerning account balances and investigate balances that differ from these expectations. In interviews dealing with the audit planning process in general, subjects made several references to having some notion of what they expect to find in an audit. Specifically, they appeared to have developed ranges within which they expected given account balance values to fall. These ranges may be imprecise and based on the expected direction of change for the account as opposed to quantified values.

Assertion 3: auditors generate expectations based on changes in events or circumstances relative to prior years. General planning interviews as well as firm documents provide evidence that determining what has changed from last year is a critical part of initial audit planning. Specific examples from data gathering sessions that involved specific clients confirmed this observation. In all cases, the main thrust of the auditor's initial data gathering during audit planning was determining what had changed from last year. The only exception is with large, sophisticated clients where the auditor has a continuous relationship with the client and becomes aware of these changes as they occur.

Assertion 4: management's incentives and abilities to manipulate account balances are of special importance in assessing inherent risk. There was ample evidence in both specific and general settings that management incentives were a key consideration in evaluating inherent risk. In addition, accounts that are subject to management judgment (e.g. the allowances for doubtful accounts and obsolete inventories) were almost always considered more risky. This emphasis on management motives and abilities is consistent with the emphasis placed by Peat Marwick International on the degree of management judgment involved in valuing accounts and the degree to which the client's financial condition affects management's incentives discussed earlier.

Additional evidence on the related issue of the importance of management incentives in choosing accounting methods appears in theoretical and empirical research in accounting. In a summary of this research, Watts & Zimmerman (1986) conclude that accounting method choices can be driven by the existence of bonus plans, bond covenants and public pressures that affect management's incentives to report certain financial and operating results.

Assertion 5: to be useful, inherent risk assessments should provide an analysis of why a given account is risky rather than merely a quantitative estimate of risk. Several quotes in the transcripts demonstrated the complexity of inherent risk assessment and the importance of interrelationships between accounts. While quantitative estimates may reflect an auditor's overall feeling about an account, they do not indicate why the account is risky or what aspects of the account contribute to its riskiness. The transcripts contain several examples of auditors making this basic point. In order to be actionable, an inherent risk assessment needs to indicate why a particular account is risky in a specific audit context.

Apart from supporting the specific assertions of the model, the transcripts indicate that they co-occur. The last three sections of Appendix includes one quote that provides a description of the overall process. The evidence found in the transcripts to support the conceptual model is further strengthened by the existence of supporting evidence from other researchers. Bouwman found evidence that financial analysts used an internal model of the firm to drive their expectations and analysis. In addition, our model is closely related to a similar model of auditing inferred by Johnson *et al.* (1987) in their analysis of auditors' fraud and error detection. The fact that they have had similar results using significantly different methods and subjects adds considerably to the validity of our conceptual model.

DEVELOPMENT OF THE COMPUTATIONAL MODEL

The main purpose of developing an initial computational model is to help determine how auditors select, evaluate and combine various factors or cues in making inherent risk assessment. The conceptual model identifies cue classes (e.g. observed changes, historical financial data, incentives) but provides only general guidance concerning how individual cues within these classes are identified as relevant to a specific case, evaluated in terms of their impact on inherent risk and combined with other cues to form a risk assessment. An initial computational model deals with each of these issues in a relatively naive way based on the initial data in order to produce behavior that an experienced auditor can evaluate.

A computational model can be presented at several different levels of abstraction. Newell (1982) describes four main levels that can be used to describe such models: device, circuit, program and knowledge. At the knowledge level, the model is viewed as an agent which is composed of a set of goals, a set of actions and a body. The agent is viewed as processing knowledge to select actions to achieve goals (Newell, 1982). Actions are synonymous with the concept of operators discussed in Newell & Simon (1972) and are responsible for affecting the agent's external environment or changing the agent's state of knowledge. The following discussion presents our computational model of inherent risk assessment as an information processing model (Newell & Simon, 1972) at the knowledge level. (More details of the program's architecture are presented in Appendix B and in Dhar *et al.*, forthcoming.) Such a presentation requires specification of the goals and actions used by the agent to solve the problem under study (Card *et al.*; 1980; Newell, 1982). Figure 2 presents an overview of the five overall goals involved in inherent risk assessment and the actions used to achieve those goals. The levels of indentation indicate a subgoal or action set designed to achieve a goal. Italicized comments provide conditional or iterative modifications of goals and/or actions.

The remainder of this section will present a more detailed description of the five key actions in Fig. 2: (1) expectation generation (which involves cue selection); (2) evidence comparison and checking; (3) incentive checking; (4) ability checking (each of which involves both cue selection and evaluation); and (5) results presentation (which involves cue combination). The remaining actions involve data access and are self-explanatory. This initial or naive model does not include an action that assesses the accounting complexity or mechanical difficulty of generating given account balances (e.g. how complex the valuation procedures or accounting standards are that relate to items in the account) because our transcripts did not provide sufficient evidence for us to determine how this assessment interfaced with other actions in the model. The following discussion deals with each of these actions in some detail.

Expectation generation

The primary assumption in expectation generation is that an account balance will reflect general historical trends unless something specific and observable occurs to alter those trends. Therefore, inputs to the computational

```
GOAL:Assess Inherent Risk - Firm Level

    GOAL:Assess Inherent Risk - Account Level
         (repeat until no more observed changes)

        GOAL:Develop Expected Values

            ACTION:Get Observed Changes
            ACTION:Get Historic Data
            ACTION:Generate Expected Value

        GOAL:Test Expectation

            ACTION:Get Actual Current Value
            ACTION:Compare expected to actual
            ACTION:Terminate Processing of Account
            (if expected and actual are close)

            GOAL:Assess Management's Incentives

                ACTION:Get Affected Incentives
                ACTION:Determine Status and Direction of Management
                       Incentives
                ACTION:Determine Management's Ability

        ACTION:Present Analysis
```

Fig. 2. Information processing model.

model consist of a series of observations about changes that have occurred in the firm or its environment. The observations contain three pieces of information:

(1) a firm related factor that has changed where a factor can be an account or some other quantifiable factor that influences that balance in a general ledger account;

(2) an estimate of the percentage change in the factor; and

(3) a brief description of the source of that observation.

For example, the auditor might observe that the client's market share should fall by 5% because a major customer of the client was purchased by a competitor of the client in a vertical merger. However, the auditor is ultimately concerned with the inherent risk of an account balance, not changes to market share. Therefore, the computational model uses a model of a firm to relate such factors as market share to sales and other accounts. Figure 3 presents a diagram of a portion of this model.

The structure of this firm model has been based on empirical evidence gathered in this study and on previous empirical studies (Bouwman, 1978). The model uses a naive assumption that this year's balance will be the same as last year's if there are no observed changes. Therefore, observed changes are processed against the prior year's balance to generate the expected balance for the current year. If there are no observed changes, the prior year's balance is used as the expectation for the current year's balance.

Evidence checking

The model next checks the expected value against the actual value in the general ledger data. The model incorporates a crude form of materiality judgment in that it uses a preset percentage range to determine whether two values are the same or not. Values that differ by less than this amount are treated as being equal. The evidence checking action drops accounts from further consideration whose expected and actual values are not materially different.

Incentive checking

The incentive checking action relates accounts that are left after the evidence checking process is completed to sets of incentives that are affected by those accounts. An incentive can

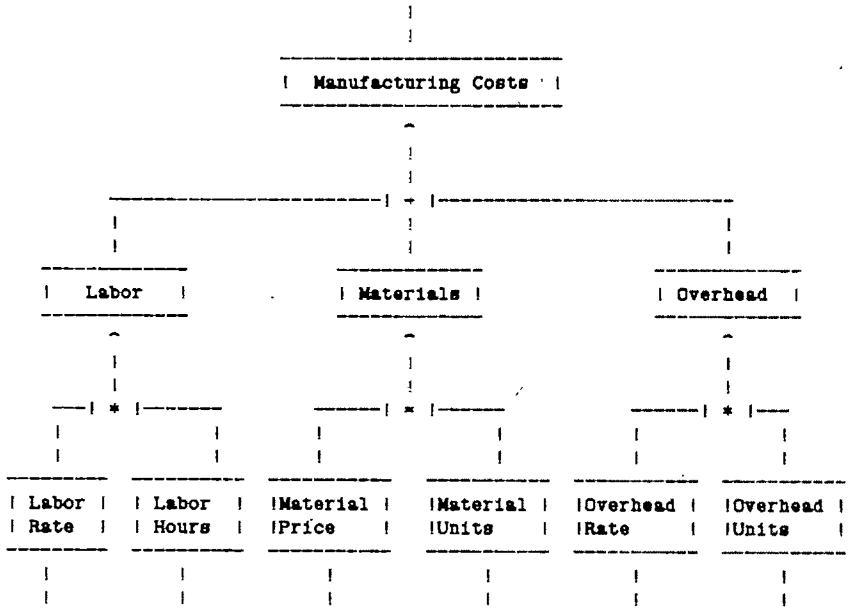


Fig. 3. Segment of the firm model.

represent either a restriction placed on account balances as part of a contract (e.g. bond covenants), contractual arrangements that tie management's compensation to levels in account balances or expectations by outsiders concerning levels of account balances (e.g. public expectations of steadily growing earnings). In general, incentives represent agreements or expectations that impact on management's motives to see that certain accounts achieve certain balances. Examples of incentives might include a management bonus plan that is triggered if gross profit exceeds 110% of last year's or a bond covenant that requires the current ratio to be greater than 2 to 1. The later incentive would be represented within the model by the formula (≥ current-ratio 2.0).

Once the model has determined the proper relationship between the general ledger account balance and the arguments in the incentive formula, it determines which direction management might be motivated to manipulate the account balance based on the current value of the incentive arguments. That is, given existence of the bonus plan mentioned earlier and given that gross profit is actually 110% of last years value, the model would determine that management might be motivated to increase sales. This determination is currently based on a very simple decision rule. The rule states that if the actual value of the incentive argument is close to the cut-off value, management might be motivated to increase the value of the argument over the cut-off by manipulating the general ledger account balance related to the argument in the appropriate direction. Given the gross profit example just cited and an actual gross profit level near 110% of last year's, the model would conclude that management might be motivated to increase gross profit by increasing sales.

If the actual value is far from the cut-off, the model assumes that management might be motivated to decrease the value of the argument in relation to the cut-off. This simple decision rule is based on the observation that most accounting decisions over which management has some control merely shift values from one year to the next. Therefore, if the argument value is either too far below a cut-off or too far above a cut-off, management might be motivated to shift values to increase the chances of making the cut-off next year. In the gross profit example, if this year's gross profit is well over 110% of last year's gross profit, management would be motivated to save some gross profit for next year by shifting sales to next year to help insure they would make the cut-off for a bonus next year. This decision rule is overly simplistic and yet captures the essence of many of the concerns expressed by auditors in the transcripts. In addition, it is consistent with empirical evidence of manager behavior on the related issue of choice of reporting methods (Healy, 1985).

Ability checking

The model incorporates a very simple ability checking procedure. Each general ledger account is coded to indicate whether or not that account is particularly sensitive to management manipulation. Such accounts as inventory and receivable valuation accounts would be coded "yes" and less sensitive accounts like cash would be coded "no". Ability checking merely involves making a note of this information for the general ledger account being processed.

Presentation of results

At this point in the processing, all the information needed to make an initial assessment of inherent risk has been accumulated. The model now pulls this information together and prints it out in an organized fashion that reflects its meaning. The current implementation of the model does not have the capability to perform any real analysis on this information such as looking for patterns in the results across accounts and drawing more general conclusions.

Test case

Appendix C presents a trace of the computational model's behavior on a test case used in its initial development. The Appendix begins by presenting an example of the data entry screen used to input the test data as well as a printout of the entire data set and concludes with a trace of the model's processing and output on that data. Italicized comments have been added to the trace to provide a more thorough explanation of the model's activities.

EVALUATION OF THE MODEL

The model has been demonstrated to several auditors who helped us design it. Their comments and critiques have been useful in helping us understand the limitations of the existing model and what must be done to alleviate them. Their general assessment was that the model was "going in the right direction" and that viewing inherent risk assessments as an expectation failure evaluation process with explicit links to management's incentives showed promise. However, they felt our initial model had several major limitations. These limitations are described below.

Overly data driven expectation generation

Our model's control of attention is driven by the observations specified by the user. However, we have now come to realize that auditors' analyses are to a large extent driven by the structure of the general ledger, beginning with current assets and ending with extraordinary items. More importantly, as this "downward scanning" proceeds, expectations about values of accounts yet to be reviewed become increasingly constrained. If the actual value does not fit with expectations, a reinterpretation of previously reviewed accounts becomes necessary. This type of reinterpretation does not take place in our model.

Overemphasis on incentives

A second criticism, related to the one above, was that our model emphasized management motives too strongly. This effect becomes even more noticeable since the model ignores the relationships among firm factors described above. While incentives are an important determinant of inherent risk, we have now found that they are usually not considered by auditors unless the value of the constrained item in the incentive formula is close to the boundary. That is, our default assumption, that account balances have been influenced by incentives, is not an accurate one. In fact, the opposite is often the case. Also, the second decision rule that shifting may occur when the value of the constrained item in the incentive formula is some distance from the bound is often applied inappropriately. This decision rule needs to be specialized. In addition, auditors did not view all incentives as having equal potential impact on management's actions. For example, violating a bond covenant would be considered more serious than not achieving a budgeted goal.

Extrapolation of data

Our model uses the prior year's balance as a basis for developing an expected balance for the current year. Auditor's tend to use more complicated processes that take into consideration recent trends in account balances. Our model will need to incorporate a trend extrapolation mechanism on which to base its expectation generation. In addition, audit planning typically occurs prior to the client's year end. Because of this, auditors do not have actual, unaudited year end balances with which to test their expectations. In order to deal with situations where only part year is available, our model will have to be expanded to annualize part year balances. In addition, some accounts do not lend themselves to extrapolation, e.g. extraordinary items. Our model will need account specific methods for creating expectations for these types of accounts.

Unintentional errors

An important addition to the current model that auditors deemed necessary was some mechanism for dealing with unintentional errors based on accounting complexity and mechanical difficulty. Examples would include miscalculating inventory values or accidentally failing to include all outstanding invoices in the accounts payable balance. Our model has no way to deal with such errors. In addition, our conceptual model does not deal with characteristics of the firm's management environment that could affect all, or at least a broad range, of accounts (e.g. management's concern for the internal control environment or high employee turnover in the accounting department). The auditors felt that these two classes of factors that influence the likelihood of an unintentional error in an account were extremely important in determining the appropriate inherent risk for a given account.

SUMMARY AND SUGGESTIONS FOR FUTURE RESEARCH

The model described here is a summary of a 2-year research effort, during which time we attempted to observe in as much detail as possible, the audit planning process with real cases. On the basis of our evidence, we have concluded that auditors do not consider it appropriate to generate numerical estimates of risk on an account by account basis, but reason about a client's financial statements using knowledge about changes in the industry and/or the client, management's motivations, prior track record and so on. Computer based models that attempt to replicate this process must be capable of modeling this range of knowledge. While our model is still inadequate in the ways described above, it is proving to be very useful in sharpening our understanding of the process of inherent risk assessment. In particular, giving auditors a real model to use (however limited) helped elicit data which would have been virtually impossible using interviewing and other data gathering techniques. Specifically, we are continuing to work with the auditors who helped us design the model as well as other auditors to determine more precisely how to enhance the model in the following ways.

(1) Provide attention directing mechanisms that follow the normal layout of a general ledger but that also recognize basic relationships between accounts and either defer judgment on a given account until all related accounts have been evaluated or reevaluate a given account when additional evidence is produced.

(2) Revise the treatment of management incentive issues to more closely match the auditor's treatment.

(3) Produce expected values based on trends in account values and account specific knowledge and provide the ability to deal with part year data.

(4) Provide mechanisms for dealing with mechanical or unintentional errors.

(5) Provide an analysis of the evidence gathered by the model that reflects patterns in the data and helps direct the auditor toward specific areas where additional data may be needed to resolve differences between expected and actual account balances.

Empirical evidence from auditors will be needed to make these enhancements to the computational model. However, the existing model provides a useful framework for soliciting the types of data needed to address these issues.

BIBLIOGRAPHY

AICPA, *AICPA Professional Standards*, Vol. A: *U.S. Auditing Standards* (AICPA, 1985).

Anderson, J. R., *The Architecture of Cognition* (Cambridge: Harvard University Press, 1983).

Anderson, J. R., Arguments Concerning Representations for Mental Imagery, *Psychological Review* (1987) pp. 249–277.

Arthur Andersen and Co., *Audit Objectives and Procedures* (1983).

Biggs, S. F. & Selfridge, M., GC-X: A Prototype Expert System for the Auditor's Going-concern Judgment, *Proceedings of the University of California Audit Judgment Symposium* (February 1986).

Boritz, J. E., Graber, B. G. & Lemon, W. M., An Experimental Study of Review of Preliminary Audit Strategy by External Auditors, *Proceedings of the Contemporary Accounting Research Conference* (August 1986).

Bouwman, M.J., Financial Diagnosis: A Cognitive Model of the Process Involved, Ph.D. dissertation, Graduate School of Industrial Administration, Carnegie-Mellon University (1978).

Card, S. K., Moran T. P. & Newell, A., Computer Text-editing: An Information-processing Analysis of a Routine Cognitive Skill, *Cognitive Psychology* (1980) pp. 32–74.

Cushing, B. E. & Loebbecke, J. K., Analytical Approaches to Audit Risk: A Survey and Analysis, *Auditing: Journal of Practice and Theory* (1983) pp. 23–41.

Dhar, V., Lewis, B. & Peters, J., A Knowledge-based Model of Audit Risk, *A. I. Magazine* (forthcoming).

Dhar, V. & Pople, H. E., Rule-based Versus Model-based Representations for Explaining and Generating Expert Behavior, *Communications of the ACM* (1987).

Dungan, C. W. & Chandler, J. S., AUDITOR: A Microcomputer-based Expert System to Support Auditors in the Field, *Expert Systems* (October 1985) pp. 210–221.

Einhorn, H. & Hogarth, R., Behavioral Decision Theory, Process of Judgment and Choice, *Annual Review of Psychology* (1981) pp. 58–88.

Elliott, R. K., Unique Audit Methods: Peat Marwick International, *Auditing: Journal of Practice and Theory* (1983) pp. 1–12.

Gibbins, M. & Wolf, F. M., Auditor's Subjective Decision Environment — The Case of a Normal External Audit, *Accounting Review* (January 1982) pp. 105–124.

Grobstein, M. & Craig, P. W., A Risk Analysis Approach to Auditing, *Auditing: A Journal of Practice and Theory* (1984) pp. 1–16.

Hansen, J. V. & Messier, W. F., Jr, A Preliminary Test of EDP-EXPERT, *Auditing: A Journal of Practice and Theory* (1986) pp. 109–123.

Healy, P. M., The Effect of Bonus Schemes on Accounting Decisions, *Journal of Accounting and Economics* (1985) pp. 85–107.

Jiambalvo, J. J. & Waller, W., Decomposition and Assessments of Audit Risk, *Auditing: Journal of Practice and Theory* (1984) pp. 80–88.

Johnson, P. E., Jamal, K. & Berryman, R. G., Audit Judgment Research: Cognition, Memory and Knowledge Representation, working paper, University of Minnesota (1987).

Libby, R., Artman, J. T. & Willingham, J. J., Process Susceptibility, Control Risk and Audit Planning, *The Accounting Review* (April 1985) pp. 212–230.

Messier, W. F., Jr & Hansen, J. V., Expert Systems in Auditing: The State of the Art, working paper, University of Florida (September 1986).

Newell, A., The Knowledge Level, *Artificial Intelligence* (1982) pp. 87–127.

Peat Marwick International, *PMI Audit Manual — U.S.* (1985).

Shpilberg, D. & Graham, L. E., Developing ExperTAX(sm): An Expert System for Corporate Tax Accrual and Planning, *Auditing: A Journal of Practice and Theory* (1986) pp. 75–94.

Stefik, M. & Bobrow, D., Object-oriented Programming: Themes and Variations, *AI Magazine* (Winter 1986) pp. 40–62.

Steinbart, P. J., Materiality: A Case Study Using Expert Systems, *Accounting Review* (January 1987) pp. 97–116.

Watts, R. L. & Zimmerman, J. L., *Positive Accounting Theory* (Prentice-Hall, 1986).

APPENDIX A. EXAMPLES OF STATEMENTS SUPPORTING MAJOR ASSERTIONS

Assessment of the account level

The following two quotes deal with auditor's reaction to a new workpaper that was being tested by one of the CPA firms involved in this research. The workpaper required auditors to rate the inherent risk of each account.

> "I think it probably lays out on paper how one's mind works; I probably think that way but I don't need that workpaper to help me think that way because I already do."
>
> "... which will take the financial statement captions and try to isolate whether we think it's high flow, low flow, medium flow, you know, what type of risk we think is associated with it in better attempt to design appropriate audit approaches and I think I'd want to do that here."

Expectation generating process

General planning interview:

> "So you almost have a preconceived notion of what you are going to see and as long as you see that you got the explanation ready for it so you don't have to do much digging."

Specific client planning meeting:

> S1 — "We, it just seems unusual with a high volume company like this you think you would develop percentages, you know, this seems like with the type of business they would be in you could develop reserve percentages but..."
>
> S2 — "Actually, you would think so but their history of write-offs has not been, has not been significant."

Differential model

General planning meeting:

> "So what I end up doing is concentrating on what's different or unique about this year versus last year. That's where you get into what's changed."

Specific client planning meeting:

> "Normally [XYZ, co.] is so simple when you think about it, debt, totally financing inventory and receivables with very little investment and fixed assets that I would have just honed in on inventory and accounts receivable, compared it to last year's audit report which we have, which I brought into the meeting and I would have said, 'well, is it level, went up or did it go down. Receivables up or down.' That type of thing."

Management's incentives and abilities

General planning interviews:

> "Why, when someone says to me, 'What is the inherent risk on the engagement?', I first think of people. That's what I do. And the strength of management and their ability of management to override what ever the systems create or to influence whatever the systems create, dramatically so."
>
> "... a sophisticated, multinational organization that's well controlled and has very formal budgetary and planning controls in its normal reporting, I think the thing that you've got to, the key factor is what is the likelihood of management overriding those controls. ... The way company management

is compensated comes into play because, to the extent that even financial management is rewarded based on the achieved level of reported results, it creates an incentive to make those reported results reach the target levels."

Need for explanations of why account is risky

General planning interviews:

"We look for interrelationships. We look to see if sales are down, you would expect receivables to be down. If sales are up and you see that receivables are down, you might ask yourself, "Has he done a better job in collecting the accounts? Did he offer a special discount program?" And if sales are up and receivables are down, has he invested the money in inventory or is he holding it in cash. You know, are all those things consistent?"

"So what's so critical about it and why. In not very general terms but very specific terms. Out of this one hundred million dollar balance what piece am I really concerned about. Okay, there's probably five million out of that one hundred million that I'm really concerned about but how do I get there?"

General support for the conceptual model

"I think the process you go through to obtain that knowledge really is to gain an understanding of the client's business, an understanding of the client, an understanding of how the fluctuations in the economy might affect a client's business. You compare the client's business to other businesses in the same industry to see if they are having consistent operating results and if not if there are logical reasons for it and if they are having consistent operating results if that is what you expected. I mean you develop expectations in your mind of what you expect to see and to the extent results don't conform to that yet. You begin asking questions and then you obtain the knowledge."

APPENDIX B. DESCRIPTION OF SYSTEM ARCHITECTURE

The primary concerns in choosing an architecture for the model were psychological validity and flexibility. Although it is difficult if not impossible to determine the psychological validity of various representation schemes at program level (Anderson, 1978), there is considerable theoretical and empirical support for separating declarative knowledge from procedural and representing declarative knowledge as a network and procedural knowledge in the form of production systems (Anderson, 1983). Therefore we elected not to use available systems based strictly on rule formalism that other researchers have used (see, for example, Dungan & Chandler, 1985; Hansen & Messier, 1986; Steinbart, 1987) and decided to develop our model directly in the LISP progamming language (see for example, Biggs & Selfridge, 1986; Shpilberg & Graham, 1986). This approach also allowed us greater flexibility (see Dhar & Pople, 1987, for a discussion of rule based systems' ability to adapt). We further elected to use an object oriented approach to add structure to the program while retaining maximum flexibility (see Stefik & Bobrow, 1986, for a discussion of object oriented programming).

Procedural knowledge

Procedural knowledge in the model was represented in four different ways: as rules, productions systems, checklists and procedures. Each representation method was implemented as a class of structured objects.

Rules. Rules primarily resided inside production systems but can also be included in checklists. Rule objects contain five fields: condition, action, variables, trace message and explanation. The condition and action fields are straightforward and represent the "if" clause and "then" clause of "if... then" rules. In keeping with the structured object formalism, rules are allowed to have local variables. However, this feature is seldom used. The trace message allows the model to display relevant information when the rule fires. This is useful for program debugging and for matching the model's process behavior against a human subject's verbal protocol. The explanation field is used like a comment line and facilitates debugging the model. Trace messages and explanation fields are common to all procedural objects and will not be discussed further.

Production systems. Production systems are rule packets that are interpreted sequentially until one fires and then the process begins at the beginning of the packet again. The processing terminates when no rules fire or when a rule tells the production system to stop. Production systems contain four fields: rule list, trace message, variables and explanation. The rule list is merely that, a list of rules that are checked sequentially until one fires. The variables represent the production system's local working memory. The values for these variables can be set by passing the production system information when it is called or by action of rules.

Checklists. Checklists represent a specialized form of production system. They execute a list of procedural objects (rules, other checklists, production systems or procedures) sequentially until all have completed their activity or one of them tells the checklist to stop. In addition, procedural objects can tell the checklist to skip other procedural objects. Checklists contains four fields: list of procedural objects, variables, trace message and explanation. The variables represent the checklist's local working memory.

Procedure. Procedures are merely nice ways for the model to call LISP procedures directly. They contain four fields: variables, trace message, explanation and procedure definition. We have attempted to minimize use of these objects in that knowledge is not as explicitly represented as it is in the other procedural objects and have limited their use to complex mathematical calculations.

Summary. The procedural objects can be linked together dynamically in call chains similar to function calls in LISP or any programming language which means the model has great flexibility. However, they are data structures that conform to certain standards of explicit knowledge representation which gives the model access to its own knowledge. This feature will be useful in future enhancements of the model that deal with improved explanation and learning.

Declarative knowledge

The declarative knowledge of the model primarily involves the firm model. The firm model is represented as a node and link network where both nodes and links are implemented as structured objects. The nodes in the network or firm model elements, are complex objects that contain a variety of fields. The most important of these are sets of formulae that determine how the values of the element can be calculated from other elements. There are also fields for various types of values (e.g. percentage change, current, prior). The formulae represent the links to other nodes and are themselves represented as structured objects. These formulae objects contain methods for calculating the various types of values associated with a firm model element.

In addition to these permanent data structures, the model utilizes temporary data structures to accumulate the results of the model's processes and to store case specific data. Data structures called expectations store the expected and actual values for a given firm model element as well as the results of the incentive and ability checking process. Complete histories are also maintained of how each value was calculated to facilitate explanation. Observation objects contain the auditor entered observations concerning explicit changes in a given firm model element. They contain three fields: change value, firm model element and a comment. The comment is not currently used by the model except to display the source of expected value calculations. Incentive objects contain formulas that indicate boundaries that management might want firm model element values to cross. Incentive objects are case specific.

The general ledger data for each case is not represented internally in the model but read in with each case from a standard ASCII file. This feature facilitates interfacing the model directly with existing data bases.

General data

The model is menu driven and there are some simple editing features to allow for easy entry of case specific data. The model reported in this paper runs on an IBM or compatible micro computer with 512K CPU and two floppy disk drives. A version currently under development requires a hard disk and 5 megabytes of CPU memory. The new version resides on a COMPAC portable II microcomputer and uses a Goldhill 286 development package. We have found that use of a portable microcomputer is a real advantage in that the model can be easily demonstrated to auditor's in their offices without any special equipment.

APPENDIX C. TRACE OF MODEL BEHAVIOR

Three types of data are entered into the model by the user: observations, incentives and general ledger account balances. General ledger balances are input from a disk file while observations and constraints are entered from a menu driven editor. The following is an example of one input screen used when the user entered an observation, **bold** *items were entered by the user.*

Adding a new observation to TEST case

Please enter requested information when prompted.
Firm or environmental element effected: **paid in capital**
Amount of expected change as a %: **5**
Source of expected change: **Public sale of common stock**
(E)nter another observation or (Q)uit? **Q**

A test case was built using this menu screen, a similar screen for incentives and a disk file for general ledger data. The following copies of model display screens present the total data set for the test case. The model displays a natural language translation of the internal data structures.

The following observations are on file:

(1) an expected 5% change in paid in capital based on public sale of common stock;
(2) an expected 7% change in common stock based on public sale of common stock;
(3) an expected −10% change in firm's market share based on purchase of customer by competitor;
(4) an expected −4% change in average manufacturing labor rate based on new union contract; and
(5) an expected −5% change in total market demand based on trade publications.

The following incentives are on file:

(1) maximum gross profit level for the management compensation package which is based on the restriction that the current year's gross profit is less than or equal to the prior year's gross profit times 1.15;

(2) minimum gross profit cut-off for the management compensation package which is based on the restriction that the current year's gross profit is greater than or equal to the prior year's gross profit times 1.1;

(3) debt to equity restriction due to a bond covenant which is based on the restriction that the current year's long term debt to equity is less than 0.5;

(4) working capital restriction due to a bond covenant which is based on the restriction that the current year's current ratio is greater than or equal to 2; and

(5) public expectation for increasing earnings which is based on the restriction that the current year's net income is greater than or equal to the prior year's net income.

The following general ledger data is on file.

Account name	Current value	Historic value
Cash	1144	1140
Short term investments	1750	1750
Accounts receivable	14,542	11,362
Ending inventory	39,364	35,388
Prepaid expenses	1087	1000
Land	1268	1268
Buildings	7114	7100
Machinery	11,109	11,000
Furniture	1674	1700
Total other assets	844	840
Current portion of long term debt	579	579
Accounts payable	15,166	15,000
Federal income taxes payable	2326	2300
Total long term debt	13,681	13,680
Convertible stock	311	311
Common stock	5581	5218
Paid in capital	7846	7499

Continued over

Continued

Account name	Current value	Historic value
Ending retained earnings	24,594	19,406
Dividends	833	833
Beginning retained earnings	19,406	12,239
Income tax expense	5724	0
Sales	121,270	123,000
Manufacturing labor	43,691	45,000
Manufacturing materials	17,477	17,000
Manufacturing overhead	26,215	26,000
Ending inventory	39,364	35,388
Beginning inventory	35,388	36,388
General and administrative expenses	26,118	26,000

The following is a trace of the system's activity given the above test data.

Generating expected change values for general ledger accounts based on 5 observations on file.

The model first propagates the change values from the 5 observations through its internal model of a firm to assess the impact of those changes on general ledger account balances.

Checking expected changes against actual change values.

Once expected changes in general ledger account balances have been calculated based on the observations and firm model, the model calculates the actual change using the general ledger data for each general ledger account.

Determining mangement motives in relation to unfulfilled expectations based on 5 constraints on file.

Determining degree of latitude for management manipulations.

Next, the model determines the relationship between the unfulfilled expectations and the list of management incentives on file.

I have completed my analysis, please hit any key to review the results.

The model concludes its analysis by presenting the data that has accumulated during the analysis. The following is a printout of all the data accumulated for one general ledger account in this test case.

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
   I did not expect a change in accounts receivable.
   The actual change in accounts receivable is 28%.
   This account is particularly sensitive to management override.
   This account affects the following constraints.
   A working capital restriction due to a bond covenant which is based on the restriction that the current
year's current ratio is greater than or equal to 2. Management may be motivated to decrease accounts re-
ceivable because current ratio is 60.2% greater than 2 which is far enough from the limit to encourage shift-
ing items to next year.
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

JOURNAL OF ACCOUNTING AND PUBLIC POLICY

Volume 7, *Number* 2 *Summer* 1988

CONTENTS

An Assessment of the Impact of Diagnosis-Related Group (DRG)-Based Reimbursement on the Technical Efficiency of New Jersey Hospitals Using Data Envelopment Analysis
JAMES P. BORDEN

Mergers and the Double Taxation of Corporate Income CHRISTOPHER R. PETRUZZI

Research Note:
Early Withdrawals from Individual Retirement Accounts (IRAs) after the 1986 Tax Reform Act
DONALD V. SAFTNER and CHERIE J. O'NEIL

Review Essay:
A Perspective on Solomons' Quest for Credibility in Financial Reporting
STEVEN B. JOHNSON

Biographies

Journal of Accounting and Public Policy is published quarterly by Elsevier Science Publishing Co., Inc., 52 Vanderbilt Avenue, New York, NY 10017. Institutional subscriptions: $125.00 per year. Personal subscriptions: $63.00 per year. Outside the United States, Canada and Mexico please add $16.00 for surface postage and handling. Single copy and back volume information available upon request.

All editorial correspondence should be sent to: Lawrence A. Gordon or Stephen E. Loeb, College of Business and Management, University of Maryland, College Park, MD 20742, U.S.A.

JOURNAL OF BUSINESS FINANCE AND ACCOUNTING

Volume 15, *Number* 3 *Autumn* 1988

Editor: Richard Briston

CONTENTS

Cash Flows, Exit Prices and British Airways A. J. ARNOLD and R. T. WEARING
UK Tax Legislation and the Switch to the Liability Method for Income Taxes THOMAS J. LINSMEIER, R. D. NAIR and JERRY J. WEYGANDT
Analytical Examination of a Business Tax Reform Under Rapid Inflation: The Israeli Case YORAM KROLL
A Comparison of Market Efficiency Among Stock Exchanges ROBERT M. BROWN
Stock Selection and Timing — A New Look at Market Efficiency GEORGE M. FRANKFURTER and CHRISTOPHER G. LAMOUREUX
Information Content of Dividend Increases: The Case of Regulated Utilities JOSEPH AHARONY, HAIM FALK and ITZHAK SWARY
An Examination of Capital Budgeting Decision Alternatives for Mutually Exclusive Investments with Unequal Lives RICHARD L. MEYER, SCOTT BELSEY and JAMES R. LONGSTREET
An Empirical Test of Operating Synergism RONALD J. KUDLA and MANJEET S. DHATT
Some Geometrical Characteristics of the Risk–Return Plane GRAHAM D. BARR and RORY F. KNIGHT
Simple Resource Allocation Rules ARUN J. PRAKASH, KRISHNAN DANDAPANI and GORDON V. KARELS

Subscriptions

	1988 volume		
	U.K. only	Overseas sterling	North America and Japan
Institutions, libraries and professional	£53.30	£66.60	$114.00
Teachers or students	£25.60	£32.00	$ 57.25

Journal of Business Finance & Accounting and *Financial Accountability & Management* (combined rate)

	U.K. only	Overseas sterling	North America and Japan
Institutions, libraries and professional	£59.00	£73.95	$125.00
Teachers or students	£27.60	£34.40	$ 58.60

Order from Journals Dept., Basil Blackwell Ltd, 108 Cowley Road, Oxford, OX4 1JF, U.K.
Editorial communications to *JBFA*/IMRAD, University of Warwick, Coventry CV4 7AL, U.K.

FINANCIAL ACCOUNTABILITY & MANAGEMENT

Volume 4, *Number* 3 *Autumn* 1988

Editor: John Perrin

CONTENTS

Accountable Management in British Central Government: Some Reflections on the Official Record
ANDREW FLYNN, ANDREW GRAY, WILLIAM I. JENKINS, BRIAN A. RUTHERFORD and WILLIAM PLOWDEN
Financial Accountability of Statutory Authorities in Australia M. MCCRAE and M. AIKEN
The Reluctant Managers: Clinicians and Budgets in the NHS
CHRISTOPHER POLLITT, STEVE HARRISON, DAVID HUNTER and GORDON MARNOCH
Assessing Competition Among Local Authorities in England and Wales PETER SMITH

Subscriptions

	1988 volume		
	U.K. only	Overseas sterling	North America and Japan
Institutions, libraries and professional	£25.50	£32.00	$52.70
Teachers or students	£11.20	£13.90	$21.00

Journal of Business Finance & Accounting and *Financial Accountability & Management* (combined rate)

Institutions, libraries and professional	£59.00	£73.95	$125.00
Teachers or students	£27.60	£34.40	$ 58.60

Order from Journals Dept., Basil Blackwell Ltd, 108 Cowley Road, Oxford, OX4 1JF, U.K.
Editorial communications to *Financial Accountability and Management*, Department of Economics, University of Exeter, Amory Building, Rennes Drive, Exeter EX4 4RJ, U.K.

Social Indicators Research

An International and Interdisciplinary Journal for Quality-of-Life Measurement

Editor

Alex C. Michalos, *Dept. of Philosophy, University of Guelph, Canada*

Since its inception in 1974, *Social Indicators Research* has emerged as the leading journal for research dealing with problems relating to the measurement of all aspects of the quality of life. These studies – empirical, philosophical, and methodological – take in the whole spectrum of society to include the individual, public and private organizations, and municipal, county, regional, national, and international systems. Topics under discussion are health, population, shelter, transportation, natural environment, social customs and morality, mental health, law enforcement, politics, education, religion, the media and the arts, science and technology, economics, poverty, and welfare.

Subscription Information **ISSN** 0303–8300
1989, Volume 21 (6 issues):
Institutional rate: Dfl. 338.00/US$165.50 *incl. p&h*
Private rate: Dfl. 134.00/US$ 56.00 *incl. p&h*

Private subscriptions should be sent direct to the publishers

Volume 4 may be ordered from:
Swets & Zeitlinger BV, P.O. Box 810, 2160 SZ LISSE, The Netherlands

Back volume information is available upon request

KLUWER
ACADEMIC
PUBLISHERS

P.O. Box 322, 3300 AH Dordrecht, The Netherlands
P.O. Box 358, Accord Station, Hingham, MA 02018-0358, U.S.A.

Accounting, Organizations and Society

AIMS & SCOPE

ACCOUNTING, ORGANIZATIONS AND SOCIETY is an international journal devoted to the behavioural, organizational and social aspects of accounting. Published material ranges from original theoretical and empirical contributions to review articles describing the state of the art in specific areas. Shorter critical assessments of experiments related to the behavioural and social aspects of accounting will also be published. Other features of the journal are book reviews and other bibliographical materials. The objectives of the journal are:

1. To provide a specialized forum for the publication of research on the behavioural, organizational and social aspects of accounting.
2. To foster new thinking, research and action on the social and behavioural aspects of accounting.
3. To report on experiments on the behavioural and social aspects of accounting in a way that explains how the experiment was developed, the process by which it was implemented and its consequences, both planned and unplanned. The journal is particularly interested in discussions of the interplay between theory, practice and social and individual values.

The journal will be concerned with all aspects of the relationship between accounting and human behaviour, organizational structures and processes, and the changing social and political environment of the enterprise. Its unique focus will cover such topics as:

- the social role of accounting, social accounting, social audit and accounting for scarce resources
- the provision of accounting information to employees and trade unions and the development of participative information systems
- processes influencing accounting innovations and the social and political aspects of accounting standard setting
- behavioural studies of the users of accounting information
- information processing views of organizations, and the relationship between accounting and other information systems and organizational structures and processes
- organizational strategies for designing accounting and information systems
- cognitive aspects of accounting and decision-making processes
- the behavioural and organizational aspects of budgeting, planning and investment appraisal

INSTRUCTIONS TO AUTHORS

The contribution of the author(s) should be an original one and it should in no way violate any existing copyright, and it should contain nothing of a libellous or scandalous nature.

Manuscripts. Three copies of the manuscript should be submitted to the Editor-in-Chief. They should be typed in double spacing on one side of the paper. A cover page should give the title of the manuscript, the author's name, position and institutional affiliation, and an acknowledgement, if desired. The title of the manuscript, but not the author's name, should appear on the first page of the text. If the manuscript refers to questionnaires of other research instruments which are not fully reproduced in the text, authors must also submit three copies of the complete research instrument.

Abstracts and Index. Three copies of an abstract not exceeding 80 words should accompany each manuscript submitted. Authors are also asked to supply a maximum of 10 "key" words or phrases which will be used for indexing purposes.

Tables and Figures. Each table and figure should have a number and brief title. The author should clearly indicate where he or she would like each table and figure placed. Each table and figure should accompany the manuscript on a separate sheet and, wherever possible, should be drawn in Indian ink on good white board, *suitable for direct reproduction*. Captions to figures should be typed consecutively on a separate page at the end of the paper.

Literature Citations. No footnotes should be used for literature citations. The work should be cited by author's name and year of publication in the body of the text, e.g. (Watson, 1975); or when reference is made to a specific page (Wilkes & Harrison, 1975, p. 21). Where the author's name is included in the text, the name should not be repeated in the reference citation: e.g. "Angrist (1975, p. 79) says . . .". For identification purposes, the suffix, *a*, *b*, etc., should follow the date when the bibliography contains more than one work published by an author in a single year.

Bibliography. The manuscript should include a bibliography containing only those references cited in the text. The references should be arranged in alphabetical order according to the surname of the first author. Works by the same author should be listed in order of publication. Each reference should contain full bibliographical details and journal titles should not be abbreviated, e.g.

King, P., Is the Emphasis of Capital Budgeting Theory Misplaced?, *Journal of Business Finance and Accounting* (Spring 1975) pp. 69–82.

Le Breton, P. P., *Administrative Intelligence — Information Systems* (Boston: Houghton Mifflin, 1969).

Footnotes. Footnotes for clarification or elaboration should be used sparingly. Where used, they should be numbered consecutively with superscript arabic numerals.

Research Instruments. Because of space limitations, questionnaires and other research instruments sometimes may not be fully reproduced in the published paper. When they are not fully reproduced a note must be inserted in the text of the paper indicating the address from which copies of the complete instrument are available.

Proofs and Copies. Proofs of papers should be checked by the author(s) and returned to the Publisher with corrections within 48 hours. At this stage, only printer's typographical errors should be corrected, and any substantial changes other than these will be charged to the author. Twenty-five copies of each paper are provided free of charge and additional copies may be ordered when proofs are returned.

Contributions: Manuscripts should be sent to the Editor-in-Chief of *ACCOUNTING, ORGANIZATIONS AND SOCIETY:*
Anthony G. Hopwood, London School of Economics and Political Science, Houghton Street, London WC2A 2AE.

Accounting, Organizations and Society

An International Journal devoted to the Behavioural, Organizational & Social Aspects of Accounting
Volume 14 Number 4 **1989**

CONTENTS

BEN-HSIEN BAO and DA-HSIEN BAO 303 LIFO adoption: a technology diffusion analysis

A. S. DUNK 321 Budget emphasis, budgetary participation and managerial performance: a note

O. A. IMOISILI 325 The role of budget data in the evaluation of managerial performance

K. KEASEY and R. WATSON 337 Consensus and accuracy in accounting studies of decision-making: a note on a new measure of consensus

L. MIA 347 The impact of participation in budgeting and job difficulty on managerial performance and work motivation: a research note

J. M. PETERS, B. L. LEWIS and V. DHAR 359 Assessing inherent risk during audit planning: the development of a knowledge based model

19 SEP 1989

Published by Pergamon Press

Indexed/abstracted in: *Curr. Cont. ASCA*,
Curr. Cont./Soc. & Beh. Sci.
and *Curr. Cont. Soc. Sci. Cit. Indx.*

Published in June 1989

AOS OPB 14 (4) 303–378 (1989)

Printed in Great Britain by Express Litho Service (Oxford)

486